

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2002, 15:51:26 ; Search time 11.65 seconds  
(without alignments)  
39.883 Million cell updates/sec

Title: US-09-727-198-5  
Perfect score: 57  
Sequence: 1 XPAPMKFFTTX 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues  
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt.40:\*  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	35	61.4	304 1 YR06_CAEEL	Q09423 caenorhabdi
2	35	61.4	327 1 DHB1_HUMAN	P14061 homo sapien
3	34	59.6	278 1 T2D7_DROME	Q27272 drosophila
4	34	59.6	453 1 TBB2_GEOCN	P32925 geotrichum
5	33	57.9	45 1 PSAT_MASLA	O31128 mastigoclad
6	33	57.9	371 1 LIG2_PHACH	P49012 phanerocha
7	33	57.9	371 1 LIG5_PHACH	P11543 phanerocha
8	33	57.9	372 1 LIG3_PHACH	P21764 phanerocha
9	33	57.9	372 1 LIG6_PHACH	P50622 phanerocha
10	33	57.9	372 1 LIG8_PHACH	P06181 phanerocha
11	33	57.9	372 1 LIGA_PHACH	P31837 phanerocha
12	33	57.9	372 1 LIGB_PHACH	P31838 phanerocha
13	33	57.9	455 1 A2AC_CAVPO	O60476 cavia porce
14	33	57.9	462 1 A2AC_HUMAN	P18823 homo sapien
15	33	57.9	476 1 PUT1_YEAST	P09368 saccharomyc
16	33	57.9	750 1 PTP2_YEAST	P29461 saccharomyc
17	33	57.9	939 1 EAE_ECO27	P19809 escherichia
18	33	57.9	969 1 MNR2_YEAST	P35724 saccharomyc
19	33	57.9	1213 1 CHS3_CANAL	P30573 candida alb
20	32	56.1	85 1 YRBA_HAETN	P45026 haemophilus
21	32	56.1	431 1 HMDH_BORBU	O51628 borrelia bu
22	32	56.1	458 1 A2AC_MOUSE	O01337 mus musculu
23	32	56.1	458 1 A2AC_MOUSE	P22086 rattus norv
24	32	56.1	499 1 VLI_HPV13	O02273 human papil
25	32	56.1	501 1 VLI_HPV55	P50820 human papil
26	32	56.1	502 1 VLI_PCPIV1	Q02274 pygmy chimp
27	32	56.1	603 1 SYD_AQUAE	O57589 aquilex aeo
28	32	56.1	730 1 PSAB_SYNY3	P29255 synechocyst
29	32	56.1	732 1 PSAB_SYNP2	P17155 synechococc
30	32	56.1	801 1 VP5_RDVA	P14583 rice dwarf
31	32	56.1	801 1 VP5_RDVF	O85437 rice dwarf
32	32	56.1	1597 1 M3K4_MOUSE	O08648 mus musculu
33	32	56.1	1607 1 M3K4_HUMAN	G95614 homo sapien

34 31.5 55.3 1337 1 YDM5\_SCHPO P87136 schizosach  
35 31 54.4 161 1 VG41\_HSVSA O01027 herpesvirus  
36 31 54.4 295 1 YPT5\_SCHPO O14210 schizosach  
37 31 54.4 344 1 DHB1\_MOUSE P51656 mus musculu  
38 31 54.4 344 1 DHB1\_MOUSE P51656 mus musculu  
39 31 54.4 382 1 CRTG\_CHLRE O57723 chlorobium  
40 31 54.4 452 1 Y397\_CLOPE P50487 clostridium  
41 31 54.4 570 1 SYG\_ARCFU O28346 archaeoglob  
42 31 54.4 622 1 FET5\_YEAST P43561 saccharomyc  
43 31 54.4 674 1 HEL5\_YTHAC O9h1x7 thermoplasm  
44 31 54.4 1325 1 G160\_MOUSE P55937 mus musculu  
45 31 54.4 1822 1 YM68\_CAEEL P34529 caenorhabdi

## ALIGNMENTS

RESULT 1  
YR06\_CAEEL STANDARD: PRT; 304 AA.  
AC 009423:  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical 34.8 kDa protein R07G3.6 in chromosome II.  
GN R07G3.6.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Connell M.;  
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U23452; AAK31546.1; -  
DR WormPep: R07G3.6; CE02025.  
KM Hypothetical protein.  
FT DOMAIN 92 96 POLY-THR.  
FT DOMAIN 132 136 POLY-THR.  
FT DOMAIN 140 144 POLY-THR.  
FT DOMAIN 211 215 POLY-PRO.  
SQ SEQUENCE 304 AA; 34846 MW; EAD74EB1BB988F05 CRC64;  
Query Match 61.4%; Score 35; DB 1; Length 304;  
Best Local Similarity 75.0%; Pred. No. 6.7;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
OY 4 PMKFFTT 11  
Db 235 PLMKFFTT 242  
RESULT 2  
DHB1\_HUMAN STANDARD: PRT; 327 AA.  
ID DHB1\_HUMAN  
AC P14061;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 15-JUN-1999 (Rel. 38, Last annotation update)  
DE Escradol 17 beta-dehydrogenase 1 (EC 1.1.1.62) (17-beta-HSD 1)  
DE (Placental 17-beta-hydroxysteroid dehydrogenase) (20 alpha-  
DE hydroxysteroid dehydrogenase) (20-alpha-HSD) (E2DH).

GN HSD17B1 OR EDH17B1 OR EDHB17 OR E17KSR.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RX MEDLINE=89384667; PubMed=2779584;  
RA Luu-The V., Labrie C., Zhao H.F., Couet J., Lachance Y., Simard J.,  
RA Leblanc G., Labrie F.;  
RT "Characterization of cDNAs for human estradiol 17 beta-dehydrogenase  
RT and assignment of the gene to chromosome 17: evidence of two mRNA  
RT species with distinct 5'-termini in human placenta.";  
RL Mol. Endocrinol. 3:1301-1309(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90231340; PubMed=2330005;  
RA Luu-The V., Labrie C., Simard J., Lachance Y., Zhao H.F., Couet J.,  
RA Leblanc G., Labrie F.;  
RT "Structure of two in tandem human 17 beta-hydroxysteroid  
RT dehydrogenase genes.";  
RL Mol. Endocrinol. 4:268-275(1990).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90328695; PubMed=2197970;  
RA Luu-The V., Labrie C., Zhao H.F., Couet J., Lachance Y., Simard J.,  
RA Cote J., Leblanc G., Lagace L., Berube D., Gagne R., Labrie F.;  
RT "Purification, cloning, complementary DNA structure, and predicted  
RT amino acid sequence of human estradiol 17 beta-dehydrogenase.";  
RL Ann. N.Y. Acad. Sci. 595:40-52(1990).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RX MEDLINE=89031223; PubMed=2846351;  
RA Pelotketo H., Isomaa V., Maenlausta O., Viikko R.;  
RT "Complete amino acid sequence of human placental 17 beta-  
RT hydroxysteroid dehydrogenase deduced from cDNA.";  
RL FEBS Lett. 239:73-77(1988).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93011163; PubMed=1327779;  
RA Pelotketo H.E., Isomaa V., Viikko R.;  
RT "Genomic organization and DNA sequences of human 17 beta-  
RT hydroxysteroid dehydrogenase genes and flanking regions. Localization  
RT of multiple Alu sequences and putative cis-acting elements.";  
RL Eur. J. Biochem. 209:459-466(1992).  
RN [6]  
RP SEQUENCE FROM N.A.  
RA Shen Y., Gibbs R.A.;  
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE OF 1-23.  
RX MEDLINE=72238162; PubMed=5045524;  
RA Burns D.J.W., Engel L.L., Bethune J.L.;  
RT "Amino acid composition and subunit structure. Human placental 17 -  
RT estradiol dehydrogenase.";  
RL Biochemistry 11:2699-2703(1972).  
RN [8]  
RP SEQUENCE OF 51-67.  
RX MEDLINE=73221216; PubMed=4719204;  
RA Nicholas J.C., Harris J.I.;  
RT "Human placental 17 -oestradiol dehydrogenase. Sequence of a tryptic  
RT peptide containing an essential cysteine.";  
RL FEBS Lett. 29:173-176(1973).  
RN [9]  
RP SEQUENCE OF 204-223.  
RX MEDLINE=86159759; PubMed=3955020;  
RA Murdock G.L., Chin C.C., Warren J.C.;  
RT "Elucidation of the chemical nature of the steady-state intermediates  
RT in the mechanism of carboxypeptidase A.";  
RL Biochemistry 25:646-651(1986).  
RN [10]

RP SEQUENCE OF 219-223.  
RX MEDLINE=84008135; PubMed=6578212;  
RA Murdock G.L., Chin C.C., Offord R.E., Bradshaw R.A., Warren J.C.;  
RT "Human placental estradiol 17 beta-dehydrogenase. Identification of a  
RT single histidine residue affinity-labeled by both 3-  
RT bromoacetoxyestron and 12 beta-bromoacetoxy-4-estrene-3,17-dione.";  
RL J. Biol. Chem. 258:11460-11464(1983).  
RN [11]  
RP SIMILARITY TO SHORT CHAIN DEHYDROGENASES.  
RX MEDLINE=89330506; PubMed=2547159;  
RA Baker M.E.;  
RT "Human placental 17 beta-hydroxysteroid dehydrogenase is homologous  
RT to Nodg protein of Rhizobium meliloti.";  
RL Mol. Endocrinol. 3:881-884(1989).  
RN [12]  
RP X-RAY CRYSTALLOGRAPHY (2.2 AND 1.7 ANGSTROMS).  
RX MEDLINE=95393021; PubMed=7663947;  
RA Chosh D., Pletnev V.Z., Zhu D.W., Wawrzak Z., Duax W.L., Pangborn W.,  
RA Labrie F., Lin S.-X.;  
RT "Structure of human estrogenic 17 beta-hydroxysteroid dehydrogenase  
RT at 2.20-A resolution.";  
RL Structure 3:503-513(1995).  
RN [13]  
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).  
RX MEDLINE=96398680; PubMed=8805577;  
RA Breton R., Housset D., Mazza C., Fontecilla-Camps J.-C.;  
RT "The structure of a complex of human 17beta-hydroxysteroid  
RT dehydrogenase with estradiol and NADP+ identifies two principal  
RT targets for the design of inhibitors.";  
RL Structure 4:905-915(1996).  
RN [14]  
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).  
RX MEDLINE=98192605; PubMed=9525918;  
RA Mazza C., Breton R., Housset D., Fontecilla-Camps J.-C.;  
RT "Unusual charge stabilization of NADP+ in 17beta-hydroxysteroid  
RT dehydrogenase.";  
RL J. Biol. Chem. 273:8145-8152(1998).  
RN [15]  
RP VARIANTS VAL-237 AND GLY-312.  
RX MEDLINE=93278403; PubMed=8389226;  
RA Normand T., Narod S., Labrie F., Simard J.;  
RT "Detection of polymorphisms in the estradiol 17 beta-hydroxysteroid  
RT dehydrogenase II gene at the EDH17B2 locus on 17q11-q21.";  
RL Hum. Mol. Genet. 2:479-483(1993).  
CC -1- FUNCTION: FAVORS THE REDUCTION OF ESTROGENS AND ANDROGENS.  
CC ALSO HAS 20-ALPHA-HSD ACTIVITY. USES PREFERENTIALLY NADH.  
CC -1- CATALYTIC ACTIVITY: Estradiol-17-beta + NAD(P)(+) = estrone +  
CC NAD(P)H.  
CC -1- PATHWAY: BIOSYNTHESIS OF ESTROGENS.  
CC -1- SUBUNIT: HOMODIMER.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
CC (SDR) FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: M27138; AAB16941.1; -  
DR EMBL: X13440; CAA31792.1; -  
DR EMBL: M36263; AAA35600.1; -  
DR EMBL: M84472; AAB16942.1; -  
DR EMBL: U34879; AAD05019.1; -  
DR PIR: A36081; DEH057.  
DR PDB: 1FDS; 12-FEB-97.  
DR PDB: 1FDU; 12-FEB-97.  
DR PDB: 1FDV; 27-MAY-98.  
DR PDB: 1FDW; 27-MAY-98.  
DR PDB: 1FDX; 27-MAY-98.



```

RESULT 4
ID TBB2_GEOCN STANDARD: PRT: 453 AA.
AC P32825;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-NOV-1993 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Tubulin beta-2 chain.
OS Geotrichum candidum (Oospora lactis).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Galactomycetes.
OX NCBI_TaxID=27317;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92079883; PubMed=1836049;
RA Gold S.E., Casale W.L., Keen N.T.;
RT "Characterization of two beta-tubulin genes from Geotrichum
candidum."
RL Mol. Genet. 230:104-112(1991).
CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S69627; AAB20557.1; ALT_SEQ.
DR PIR: S18597; S18597.
DR InterPro: IPR002453; Beta_tubulin.
DR InterPro: IPR000217; Tubulin.
DR InterPro: IPR003008; Tubulin_ftsz.
DR Pfam: PF000091; tubulin; 1.
DR PRINTS: PR01161; TUBULIN.
DR PROSITE: PS00227; TUBULIN.
DR PROSITE: PS00228; TUBULIN_B_AUTOREG; 1.
DR Microtubules; GTP-binding; Multigene family.
KW NP_BIND 142 148
FT
SQ SEQUENCE 453 AA; 50399 MW; 2EA9D8A0246E0371 CRC64;

Query Match 59.6%; Score 34; DB 1; Length 453;
Best Local Similarity 60.0%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 PAPMKFFTT 11
DB 261 PPRRLHFTT 270

RESULT 5
ID PSAL_MASLA STANDARD: PRT: 45 AA.
AC O31128;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Photosystem I reaction center subunit IX.
GN PSAL.
OS Mastigocladus laminosus (Fischerella sp.).
OC Bacteria; Cyanobacteria; Stigonematales; Fischerella.
OX NCBI_TaxID=1191;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=PCC 7605;

```

```

RA He 2.-Y., Chitnis V.P., Chitnis P.R., Nechustai R.;
RT "Molecular cloning of the psaf and psal genes of photosystem I from
RT the thermophilic cyanobacterium Mastigocladus laminosus."
RL (In) Plant Gene Register PGR98-026.
CC -1- FUNCTION: May help in the organization of the psae and psaf
CC subunits (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PSAL FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF030004; AAC04843.1; -.
DR InterPro: IPR002615; PSI_Psal.
DR Pfam: PF01701; PSI_Psal; 1.
DR ProDom: PD004198; PSI_Psal; 1.
KW Photosystem I; Photosynthesis; Transmembrane.
FT TRANSMEM 10 30
SQ SEQUENCE 45 AA; 5202 MW; 9B33538512A02437 CRC64;

Query Match 57.9%; Score 33; DB 1; Length 45;
Best Local Similarity 55.6%; Pred. No. 2;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 3 APMKFFTT 11
DB 5 SPLKFLTT 13

RESULT 6
ID LG2_PHACH STANDARD: PRT: 371 AA.
AC P49012;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ligninase LG2 precursor (EC 1.11.1.-) (Lignin peroxidase).
GN LG2 OR LIP2.
OS Phanerochaete chrysosporium.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Aphyllophorales; Corticiaceae; Phanerochaete.
OX NCBI_TaxID=5306;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-48.
RX STRAIN=OGC101;
RX MEDLINE=92077421; PubMed=1743510;
RA Ritch T.G., Jr., Nipper V.J., Akleewaran L., Smith A.J., Prihnow D.G.,
RA Gold M.H.;
RT "Lignin peroxidase from the basidiomycete Phanerochaete chrysosporium
RT is synthesized as a preproenzyme."
RL Gene 107:119-126(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=OGC101;
RX MEDLINE=92380494; PubMed=1511887;
RA Ritch T.G., Jr., Gold M.H.;
RT "Characterization of a highly expressed lignin peroxidase-encoding
RT gene from the basidiomycete Phanerochaete chrysosporium."
RL Gene 118:73-80(1992).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).
RX STRAIN=ATCC 24725 / BKM 1767;
RX MEDLINE=99150427; PubMed=10024453;
RA Choinowski T., Blodig W., Winterhalter K.H., Piontek K.;
RT "The crystal structure of lignin peroxidase at 1.70-A resolution
RT reveals a hydroxy group on the beta of tryptophan 171: a novel
RT radical site formed during the redox cycle."
RL J. Mol. Biol. 286:809-827(1999).

```

RT	"Genomic organization of lignin peroxidase genes of Phanerochaete									
RT	chrysosporium.";									
RL	Nucleic Acids Res. 19:599-603(1991).									
RN	[3]									
RP	SEQUENCE FROM N.A.									
RC	STRAIN-ATCC 24725 / BKM 1767;									
RX	MEDLINE-91153647; PubMed-1999283;									
RA	Zhang Y.-Z., Reddy C.A., Rasooly A.;									
RT	"Cloning of several lignin peroxidase (LIP)-encoding genes: sequence									
RT	analysis of the LIP6 gene from the white-rot basidiomycete,									
RT	Phanerochaete chrysosporium.";									
RL	Gene 97:191-198(1991).									
CC	-1- CATALYTIC ACTIVITY: DEPOLYMERIZATION OF LIGNIN. CATALYSES THE									
CC	C(ALPHA)-C(BETA) CLEAVAGE OF THE PROPYL SIDE CHAINS OF LIGNIN.									
CC	-1- PATHWAY: FIRST STEP IN LIGNIN DEGRADATION.									
CC	-1- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. LIGNINASE SUPERFAMILY.									
CC	-----									
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration									
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -									
CC	the European Bioinformatics Institute. There are no restrictions on its									
CC	use by non-profit institutions as long as its content is in no way									
CC	modified and this statement is not removed. Usage by and for commercial									
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>									
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).									
CC	-----									
DR	EMBL; M18794; AAA33734.1; -;									
DR	EMBL; X55343; CAA39033.1; -;									
DR	EMBL; M63496; AAA33739.1; -;									
DR	PIR; JN0117; OPIGGS.									
DR	HSSP; P06181; 1B85.									
DR	InterPro: IPR002016; Peroxidase.									
DR	Pfam; PF00141; Peroxidase; 1.									
DR	PRINTS; PR00458; PEROXIDASE.									
DR	PROSITE; PS00435; PEROXIDASE_1; 1.									
DR	PROSITE; PS00436; PEROXIDASE_2; 1.									
KW	Oxidoreductase; Peroxidase; Heme; Glycoprotein; Multigene family;									
KW	Lignin degradation; Signal; Zymogen.									
FT	SIGNAL 1 21									
FT	PROPEP 22 27									
FT	CHAIN 28 371									
FT	ACT_SITE 70 70									
FT	ACT_SITE 74 74									
FT	ACT_SITE 202 202									
FT	DISULFID 30 42									
FT	DISULFID 41 311									
FT	DISULFID 61 146									
FT	DISULFID 275 344									
FT	CARBOHYD 283 283									
SO	SEQUENCE 371 AA; 39417 MW; F0AE30C69D7EF5C2 CRC64;									
/										
Query Match	57.9%; Score 33; DB 1; Length 371;									
Best Local Similarity	75.0%; Pred. No. 21;									
Matches 6;	Conservative 0; Mismatches 2; Indels 0; Gaps 0;									
QY	3 APMKFFT 10									
DB	149 APQNFET 156									
RESULT 8										
LIG3	LIG3_PHACH STANDARD; PRT; 372 AA.									
AC	P21764;									
DT	01-MAY-1991 (Rel. 18, Created)									
DT	01-MAY-1991 (Rel. 18, last sequence update)									
DT	16-OCT-2001 (Rel. 40, last annotation update)									
DE	Ligninase Ig3 precursor (EC 1.11.1.-) (Lignin peroxidase).									
GN	GIG3 OR LIP.									
OS	Phanerochaete chrysosporium.									
CC	Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;									
CC	Ascomycota; Basidiomycota; Corticiaceae; Phanerochaete.									
NCBI	TaxID=5306;									

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 24725 / BKM 1767;
RA MEDLINE=91088334; PubMed=2129560;
RX Naidu P.S., Reddy C.A.;
RT "Nucleotide sequence of a new lignin peroxidase gene GLG3 from the
RL white-rot fungus, Phanerochaete chrysosporium.";
RL Nucleic Acids Res. 18:7173-7173(1990).
CC -i- CATALYTIC ACTIVITY: DEPOLYMERIZATION OF LIGNIN. CATALYZES THE
CC C(ALPHA)-C(BETA) CLEAVAGE OF THE PROPYL SIDE CHAINS OF LIGNIN.
CC -i- PATHWAY: FIRST STEP IN LIGNIN DEGRADATION.
CC -i- DEVELOPMENTAL STAGE: LIGNINASES ARE EXPRESSED DURING SECONDARY
CC METABOLISM, AND ARE TRIGGERED BY NUTRIENT LIMITATION.
CC -i- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. LIGNINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X51590; CAA35939.1; -.
DR PIR: S13723; OPJG3P.
DR HSSP: P06181; 1B80.
DR InterPro: IPR002016; Peroxidase.
DR Pfam: PF00141; peroxidase; 1.
DR PRINTS: PR00458; PEROXIDASE.
DR PROSITE: PS00435; PEROXIDASE_1; 1.
DR PROSITE: PS00436; PEROXIDASE_2; 1.
DR Oxidoreductase; Peroxidase; Heme; Glycoprotein; Multigene family;
KW Lignin degradation; Signal; Zymogen.
FT SIGNAL 1 21
FT PROPEP 22 28
FT CHAIN 29 372
FT ACT_SITE 71 71
FT ACT_SITE 75 75
FT DISULFID 31 43
FT DISULFID 62 148
FT DISULFID 277 345
FT CARBOHYD 285 285
SQ SEQUENCE 372 AA; 39536 MW; DA0293178F4E2A4A CRC64;

Query Match 57.9%; Score 33; DB 1; Length 372;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 APMKFEFT 10
   |||
Db 151 APQMNFEFT 158

RESULT 9
LIG6_PHACH STANDARD; PRT; 372 AA.
AC P50622;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ligninase L66 precursor (EC 1.11.1.-) (Lignin peroxidase).
GN GLG6.
OS Phanerochaete chrysosporium.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Aphyllophorales; Corticiaceae; Phanerochaete.
OX NCBI_TaxID=5306;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 24725 / BKM 1767;
RX MEDLINE=91097609; PubMed=2268360;
RA Naidu P.S., Zhang Y.Z., Reddy C.A.;

```

```

RT "Characterization of a new lignin peroxidase gene (GLG6) from
RT Phanerochaete chrysosporium.";
RL Biochem. Biophys. Res. Commun. 173:994-1000(1990).
CC -i- CATALYTIC ACTIVITY: DEPOLYMERIZATION OF LIGNIN. CATALYZES THE
CC C(ALPHA)-C(BETA) CLEAVAGE OF THE PROPYL SIDE CHAINS OF LIGNIN.
CC -i- PATHWAY: FIRST STEP IN LIGNIN DEGRADATION.
CC -i- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. LIGNINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M80213; AAA3737.1; -.
DR HSSP: P06181; 1B80.
DR InterPro: IPR002016; Peroxidase.
DR Pfam: PF00141; peroxidase; 1.
DR PRINTS: PR00458; PEROXIDASE.
DR PROSITE: PS00435; PEROXIDASE_1; 1.
DR PROSITE: PS00436; PEROXIDASE_2; 1.
DR Oxidoreductase; Peroxidase; Heme; Glycoprotein; Multigene family;
KW Lignin degradation; Signal; Zymogen.
FT SIGNAL 1 21
FT PROPEP 22 28
FT CHAIN 29 372
FT ACT_SITE 71 71
FT ACT_SITE 75 75
FT DISULFID 31 43
FT DISULFID 62 148
FT DISULFID 277 345
FT CARBOHYD 285 285
SQ SEQUENCE 372 AA; 39259 MW; B192A93351F8CC16 CRC64;

Query Match 57.9%; Score 33; DB 1; Length 372;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 APMKFEFT 10
   |||
Db 151 APQMNFEFT 158

RESULT 10
LIG8_PHACH STANDARD; PRT; 372 AA.
AC P06181;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ligninase H8 precursor (EC 1.11.1.-) (Lignin peroxidase).
GN LPOA.
OS Phanerochaete chrysosporium.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Aphyllophorales; Corticiaceae; Phanerochaete.
OX NCBI_TaxID=5306;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 24725 / BKM 1767;
RX MEDLINE=87173020; PubMed=3561490;
RA Tien M., Tu C.-P.D.;
RT "Cloning and sequencing of a cDNA for a ligninase from Phanerochaete
RT chrysosporium.";
RL Nature 326:520-523(1987).
RN [2]
RP REVISIONS.
RA Tien M., Tu C.-P.D.;
RL Nature 328:742-742(1987).
RN [3]

```

RP SEQUENCE FROM N.A.  
 RX MEDLINE=89196904; PubMed=240864;  
 RA Walther L., Kaelin M., Reiser J., Suter F., Fritsche B.,  
 RA Saloheimo M., Leisola M., Teeri T., Knowles J.K.C., Fiechter A.;  
 RT "Molecular analysis of a Phanerochaete chrysosporium lignin  
 peroxidase gene.";  
 RL Gene 70:127-137(1988).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88144011; PubMed=3344218;  
 RA Smith T.L., Schaich H., Gaskell J., Covert S., Cullian D.;  
 RT "Nucleotide sequence of a ligninase gene from Phanerochaete  
 chrysosporium.";  
 RL Nucleic Acids Res. 16:1219-1219(1988).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89334875; PubMed=2474293;  
 RA Andrawis A., Pease E.A., Kuan I., Holzdaur E., Tien M.;  
 RT "Characterization of two lignin peroxidase clones from Phanerochaete  
 chrysosporium.";  
 RL Biochem. Biophys. Res. Commun. 162:673-680(1989).  
 RN [6]  
 RP SEQUENCE OF 1-1 FROM N.A.  
 RX MEDLINE=88339943; PubMed=2844176;  
 RA Holzdaur E.L.F., Tien M.;  
 RT "Structure and regulation of a lignin peroxidase gene from  
 Phanerochaete chrysosporium.";  
 RL Biochem. Biophys. Res. Commun. 155:626-633(1988).  
 RN [7]  
 RP SEQUENCE OF 29-58.  
 RX STRAIN-ATCC 24725 / BKM 1767;  
 RC MEDLINE=90151655; PubMed=2303054;  
 RA Glumoff T., Harvey P.J., Molinari S., Goble M., Frank G., Palmer J.M.,  
 RA Smilt J.D.G., Leisola M.S.A.;  
 RT "Lignin peroxidase from Phanerochaete chrysosporium. Molecular and  
 kinetic characterization of isozymes.";  
 RL Eur. J. Biochem. 187:515-520(1990).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).  
 RC STRAIN-ATCC 24725 / BKM 1767;  
 RX MEDLINE=98301364; PubMed=9636023;  
 RA Blodig W., Doyle W.A., Smith A.T., Winterhalter K., Choinowski T.,  
 RA Plonczek K.;  
 RT "Autocatalytic formation of a hydroxy group at C beta of Trp171 in  
 lignin peroxidase.";  
 RL Biochemistry 37:8832-8838(1998).  
 CC -1- CATALYTIC ACTIVITY: DEPOLYMERIZATION OF LIGNIN, CATALYSES THE  
 C CATHA-C(BETA) CLEAVAGE OF THE PROPYL SIDE CHAINS OF LIGNIN.  
 CC -1- PATHWAY: FIRST STEP IN LIGNIN DEGRADATION.  
 CC -1- DEVELOPMENTAL STAGE: LIGNINASES ARE EXPRESSED DURING SECONDARY  
 CC METABOLISM, AND ARE TRIGGERED BY NUTRIENT LIMITATION.  
 CC -1- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY, LIGNINASE SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC  
 DR EMBL; M37701; AAA33740.1; ALT\_SEQ.  
 DR EMBL; Y00262; CAA68373.1; ALT\_SEQ.  
 DR EMBL; X06689; CAA29882.1; -  
 DR EMBL; M21913; -; NOT ANNOTATED\_CDS.  
 DR EMBL; M27401; AAA53109.1; -  
 DR EMBL; M27884; AAB00798.1; -  
 DR PIR; A27817; A27817.  
 DR PIR; S01028; S01028.  
 DR PIR; J04042; J04042.  
 DR PIR; S08202; S08202.  
 DR PDB; 1B80; 1B-MAR-99.  
 DR PDB; 1B82; 05-MAR-99.

DR PDB; 1B85; 16-FEB-99.  
 DR InterPro; IPR002016; Peroxidase.  
 DR Pfam; PF00141; peroxidase\_1.  
 DR PRINTS; PR00458; PEROXIDASE.  
 DR PROSITE; PS00435; PEROXIDASE\_1; 1.  
 DR PROSITE; PS00436; PEROXIDASE\_2; 1.  
 KW Oxidoreductase; Peroxidase; Heme; glycoprotein; Multigene family;  
 KW Lignin degradation; Signal; Zymogen; 3D-structure.  
 FT SIGNAL 1 21  
 FT PROPEP 22 28  
 FT CHAIN 29 372  
 FT ACT\_SITE 71 71  
 FT ACT\_SITE 75 75  
 FT ACT\_SITE 204 204  
 FT DISULFID 31 43  
 FT DISULFID 42 313  
 FT DISULFID 62 148  
 FT DISULFID 277 345  
 FT CARBOHYD 285 285  
 FT VARIANT 15 15  
 FT VARIANT 142 142  
 FT SEQUENCE 372 AA; 39685 MW; E6ABA9FD48428FCC CRC64;  
 SO  
 Query Match 57.9%; Score 33; DB 1; Length 372;  
 Best Local Similarity 75.0%; Pred. No. 21;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 APMAKFFT 10  
 DB 151 APOMNFFT 158  
 ID LIGA\_PNHCH STANDARD: PRT; 372 AA.  
 AC P31837;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Ligninase A precursor (EC 1.11.1.-) (Lignin peroxidase).  
 GN LIPA OR LPOB.  
 OS Phanerochaete chrysosporium.  
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;  
 OC Aphyllophorales; Corticiaceae; Phanerochaete.  
 OX NCBI\_TaxID=5306;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN-ATCC 24725 / BKM 1767;  
 RC MEDLINE=91187681; PubMed=2011531;  
 RA Gaskell J., Dieperink E., Cullen D.;  
 RT "Genomic organization of lignin peroxidase genes of Phanerochaete  
 chrysosporium.";  
 RL Nucleic Acids Res. 19:599-603(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN-ATCC 24725 / BKM 1767;  
 RC MEDLINE=90323600; PubMed=2373364;  
 RA Huoponen K., Ollikka P., Kaelin M., Walther I., Meentsaelae P.,  
 RA Reiser J.;  
 RT "Characterization of lignin peroxidase-encoding genes from lignin-  
 RT degrading basidiomycetes.";  
 RL Gene 89:145-150(1990).  
 CC -1- CATALYTIC ACTIVITY: DEPOLYMERIZATION OF LIGNIN, CATALYSES THE  
 CC CATHA-C(BETA) CLEAVAGE OF THE PROPYL SIDE CHAINS OF LIGNIN.  
 CC -1- PATHWAY: FIRST STEP IN LIGNIN DEGRADATION.  
 CC -1- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY, LIGNINASE SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; X54257; CAA38177.1; -.
CC PIR; S13564; OPUGBP.
CC HSSP; P06181; 1882.
CC InterPro; IPR002016; Peroxidase.
DR EMBL; M37701; AAA33741.1; -.
DR PIR; S13563; OPUGAP.
DR HSSP; P49012; ILIP.
DR InterPro; IPR002016; Peroxidase.
DR Pfam; PF00141; peroxidase; 1.
DR PRINTS; PR00458; PEROXIDASE.
DR PROSITE; PS00435; PEROXIDASE_2; 1.
DR PROSITE; PS00436; PEROXIDASE_1; 1.
DR Oxidoreductase; Peroxidase; Heme; Glycoprotein; Multigene family;
KW Lignin degradation; Signal; Zymogen.
FT SIGNAL 1 21
FT PROPE 22 28
FT CHAIN 29 372
FT ACT_SITE 71 71
FT ACT_SITE 75 75
FT ACT_SITE 204 204
FT DISULFID 31 43
FT DISULFID 42 313
FT DISULFID 62 148
FT DISULFID 277 345
FT CAROHND 285 285
FT CONFLICT 135 135
FT CONFLICT 161 161
FT CONFLICT 298 298
SQ SEQUENCE 372 AA; 39393 MW; 0ACE4598D6D6F09B CRC64;

Query Match 57.9%; Score 33; DB 1; Length 372;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 APMKRFET 10
DB 151 APOMNFT 158

RESULT 12
LIBG PHACH STANDARD; PRT; 372 AA.
AC P31838;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ligninase B precursor (EC 1.11.1.1) (Lignin peroxidase).
GN LIPB.
OS Phanerochaete chrysosporium.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Aphyllophorales; Corticiaceae; Phanerochaete.
OX NCBI_TaxId=5306;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 24725 / BKM 1767;
RA MEDLINE=91187681; PubMed=2011531;
RA Gaskell J., Diaperink E., Cullen D.;
RT "Genomic organization of lignin peroxidase genes of Phanerochaete
chrysosporium";
RL Nucleic Acids Res. 19:599-603(1991).
CC -1- CATALYTIC ACTIVITY: DEPOLYMERIZATION OF LIGNIN. CATALYSES THE
C(ALPHA)-C(BETA) CLEAVAGE OF THE PROPYL SIDE CHAINS OF LIGNIN.
CC -1- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. LIGNINASE SUBFAMILY.
CC -1- PATHWAY: FIRST STEP IN LIGNIN DEGRADATION.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

```

```

CC EMBL; X54257; CAA38178.1; -.
CC PIR; S13564; OPUGBP.
CC HSSP; P06181; 1882.
CC InterPro; IPR002016; Peroxidase.
DR EMBL; M37701; AAA33741.1; -.
DR PIR; S13563; OPUGAP.
DR HSSP; P49012; ILIP.
DR InterPro; IPR002016; Peroxidase.
DR Pfam; PF00141; peroxidase; 1.
DR PRINTS; PR00458; PEROXIDASE.
DR PROSITE; PS00435; PEROXIDASE_1; 1.
DR PROSITE; PS00436; PEROXIDASE_2; 1.
KW Oxidoreductase; Peroxidase; Heme; Glycoprotein; Multigene family;
KW Lignin degradation; Signal; Zymogen.
FT SIGNAL 1 21
FT PROPE 22 28
FT CHAIN 29 372
FT ACT_SITE 71 71
FT ACT_SITE 75 75
FT ACT_SITE 204 204
FT DISULFID 31 43
FT DISULFID 42 313
FT DISULFID 62 148
FT DISULFID 277 345
FT CAROHND 285 285
SQ SEQUENCE 372 AA; 39530 MW; E0DD2CF40425D9D5 CRC64;

Query Match 57.9%; Score 33; DB 1; Length 372;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 APMKRFET 10
DB 151 APOMNFT 158

RESULT 13
A2AC CAVPO STANDARD; PRT; 455 AA.
AC 060476;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Alpha-2C adrenoceptor receptor (Alpha-2C adrenoceptor).
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriognathii; Cavidae; Cavia.
OX NCBI_TaxId=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HARTLEY.
RA MEDLINE=96152573; PubMed=8573196;
RA Svensson S.P., Bailey T.J., Porter A.C., Richman J.G., Regan J.W.;
RT "Heterologous expression of the cloned guinea pig alpha 2A, alpha 2B,
and alpha 2C adrenoceptor subtypes. Radioligand binding and
functional coupling to a CAMP-responsive reporter gene.";
RL Biochem. Pharmacol. 51:291-300(1996).
CC -1- FUNCTION: ALPHA-2 ADRENERGIC RECEPTORS MEDIATE THE CATECHOLAMINE-
INDUCED INHIBITION OF ADENYLATE CYCLASE THROUGH THE ACTION OF G
PROTEINS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL; U25724; AAA67076.1; -.
CC InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.

```



DR PROSITE; PS00237; G-PROTEIN\_RECEP\_FL-1; 1.  
 DR PROSITE; PS50262; G-PROTEIN\_RECEP\_FL-2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KW Multigene family; Phosphorylation.  
 FT DOMAIN 1 47  
 FT TRANSSEM 48 72  
 FT TRANSSEM 73 84  
 FT TRANSSEM 85 110  
 FT TRANSSEM 111 120  
 FT TRANSSEM 121 143  
 FT TRANSSEM 144 164  
 FT TRANSSEM 165 187  
 FT TRANSSEM 188 203  
 FT TRANSSEM 204 227  
 FT TRANSSEM 377 400  
 FT TRANSSEM 401 413  
 FT TRANSSEM 414 434  
 FT TRANSSEM 435 455  
 FT CARBOHYD 17 17  
 FT CARBOHYD 29 29  
 FT DISULFID 120 198  
 FT DISULFID 120 198  
 SO SEQUENCE 455 AA; 49351 MW; 6B657D247FF8A3F4 CRC64;

Query Match 57.98; Score 33; DB 1; Length 455;  
 Best Local Similarity 62.5%; Pred. No. 27;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 PAPMKPF 9  
 Db 408 PTPLEKFF 415

RESULT 14  
 A2AC\_HUMAN STANDARD; PRT; 462 AA.  
 AC P18825; P35369; Q9HB49;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-MAR-2002 (Rel. 41, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Alpha-2C-adrenergic receptor (Alpha-2C adrenoceptor) (Subtype C4).  
 GN ADRA2C.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 ON NCBI\_TaxID=9606;  
 RX SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE-Kidney;  
 RX MEDLINE=88320430; PubMed=2842764;  
 RA Regan J.W., Koblika T.S., Yang-Feng T.L., Caron M.G., Lefkowitz R.J.,  
 RA Koblika B.K.;  
 RT "Cloning and expression of a human kidney cDNA for an alpha  
 RT 2-adrenergic receptor subtype."  
 RL Proc. Natl. Acad. Sci. U.S.A. 85:6301-6305(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE-Brain;  
 RX MEDLINE=98041882; PubMed=9371698;  
 RA Schack S., Devedjian J.C., Cayla C., Sender Y., Paris H.;  
 RT "Molecular cloning, sequencing and functional study of the promoter  
 RT region of the human alpha2C4-adrenergic receptor gene."  
 RL Biochem. J. 328:431-438(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RA Yano K., Takeda M., Sugimoto E., Sagai H.;  
 RT "Molecular cloning and expression of a novel human alpha2C-adrenergic  
 RT receptor, alpha2CII, gene."  
 RL Submitted (OCT-1992) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT 322-GLY--PRO-325 DEL.  
 RX MEDLINE=20390061; PubMed=10801795;

RA Small K.M., Forbes S.L., Rahman F.F., Bridges K.M., Liggett S.B.;  
 RT "A four amino acid deletion polymorphism in the third intracellular  
 RT loop of the human alpha 2C-adrenergic receptor confers impaired  
 RT coupling to multiple effectors."  
 RL J. Biol. Chem. 275:23059-23064(2000).  
 CC -1- FUNCTION: ALPHA-2 ADRENERGIC RECEPTORS MEDATE THE CATECHOLAMINE-  
 CC INDUCED INHIBITION OF ADENYLYLATE CYCLASE THROUGH THE ACTION OF G  
 CC PROTEINS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are  
 CC produced by alternative splicing.  
 CC -1- POLYMORPHISM: The Del322-325 variant has a significant loss of  
 CC function. It is approximately 10 times more frequent in African-  
 CC Americans compared with Caucasians (allele frequencies 0.381  
 CC versus 0.040).  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; J03853; AAA5513.1; -;  
 CC EMBL; U72648; AAC78723.1; -;  
 CC EMBL; D13538; BAA02737.1; -;  
 CC EMBL; AF280399; AAC628076.1; -;  
 CC EMBL; AF280400; AAC628077.1; -;  
 CC PIR; A31237; A31237.  
 CC HSSP; P29274; IMM.  
 CC GCRDB; GCR\_0044; -;  
 CC GCRDB; GCR\_0484; -;  
 CC GCRDB; GCR\_1903; -;  
 CC MIM; 104250; -;  
 CC InterPro; IPR000276; GPCR\_Rhodopsn.  
 CC Pfam; PF00001; 7tm\_1; 1.  
 CC PRINTS; PR00237; GPCR\_RHODOPSIN.  
 DR PROSITE; PS00237; G-PROTEIN\_RECEP\_FL-1; 1.  
 DR PROSITE; PS50262; G-PROTEIN\_RECEP\_FL-2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KW Multigene family; Phosphorylation; Polymorphism; Alternative splicing.  
 FT DOMAIN 1 51  
 FT TRANSSEM 52 76  
 FT TRANSSEM 77 88  
 FT TRANSSEM 89 114  
 FT TRANSSEM 115 124  
 FT TRANSSEM 125 147  
 FT TRANSSEM 148 168  
 FT TRANSSEM 169 191  
 FT TRANSSEM 192 207  
 FT TRANSSEM 208 231  
 FT TRANSSEM 232 379  
 FT TRANSSEM 380 407  
 FT TRANSSEM 408 420  
 FT TRANSSEM 421 441  
 FT TRANSSEM 442 462  
 FT TRANSSEM 462 484  
 FT TRANSSEM 484 506  
 FT TRANSSEM 507 529  
 FT TRANSSEM 530 552  
 FT TRANSSEM 553 575  
 FT TRANSSEM 576 598  
 FT TRANSSEM 599 621  
 FT TRANSSEM 622 644  
 FT TRANSSEM 645 667  
 FT TRANSSEM 668 690  
 FT TRANSSEM 691 713  
 FT TRANSSEM 714 736  
 FT TRANSSEM 737 759  
 FT TRANSSEM 760 782  
 FT TRANSSEM 783 805  
 FT TRANSSEM 806 828  
 FT TRANSSEM 829 851  
 FT TRANSSEM 852 874  
 FT TRANSSEM 875 897  
 FT TRANSSEM 898 920  
 FT TRANSSEM 921 943  
 FT TRANSSEM 944 966  
 FT TRANSSEM 967 989  
 FT TRANSSEM 990 1012  
 FT TRANSSEM 1013 1035  
 FT TRANSSEM 1036 1058  
 FT TRANSSEM 1059 1081  
 FT TRANSSEM 1082 1104  
 FT TRANSSEM 1105 1127  
 FT TRANSSEM 1128 1150  
 FT TRANSSEM 1151 1173  
 FT TRANSSEM 1174 1196  
 FT TRANSSEM 1197 1219  
 FT TRANSSEM 1220 1242  
 FT TRANSSEM 1243 1265  
 FT TRANSSEM 1266 1288  
 FT TRANSSEM 1289 1311  
 FT TRANSSEM 1312 1334  
 FT TRANSSEM 1335 1357  
 FT TRANSSEM 1358 1380  
 FT TRANSSEM 1381 1403  
 FT TRANSSEM 1404 1426  
 FT TRANSSEM 1427 1449  
 FT TRANSSEM 1450 1472  
 FT TRANSSEM 1473 1495  
 FT TRANSSEM 1496 1518  
 FT TRANSSEM 1519 1541  
 FT TRANSSEM 1542 1564  
 FT TRANSSEM 1565 1587  
 FT TRANSSEM 1588 1610  
 FT TRANSSEM 1611 1633  
 FT TRANSSEM 1634 1656  
 FT TRANSSEM 1657 1679  
 FT TRANSSEM 1680 1702  
 FT TRANSSEM 1703 1725  
 FT TRANSSEM 1726 1748  
 FT TRANSSEM 1749 1771  
 FT TRANSSEM 1772 1794  
 FT TRANSSEM 1795 1817  
 FT TRANSSEM 1818 1840  
 FT TRANSSEM 1841 1863  
 FT TRANSSEM 1864 1886  
 FT TRANSSEM 1887 1909  
 FT TRANSSEM 1910 1932  
 FT TRANSSEM 1933 1955  
 FT TRANSSEM 1956 1978  
 FT TRANSSEM 1979 2001  
 FT TRANSSEM 2002 2024  
 FT TRANSSEM 2025 2047  
 FT TRANSSEM 2048 2070  
 FT TRANSSEM 2071 2093  
 FT TRANSSEM 2094 2116  
 FT TRANSSEM 2117 2139  
 FT TRANSSEM 2140 2162  
 FT TRANSSEM 2163 2185  
 FT TRANSSEM 2186 2208  
 FT TRANSSEM 2209 2231  
 FT TRANSSEM 2232 2254  
 FT TRANSSEM 2255 2277  
 FT TRANSSEM 2278 2300  
 FT TRANSSEM 2301 2323  
 FT TRANSSEM 2324 2346  
 FT TRANSSEM 2347 2369  
 FT TRANSSEM 2370 2392  
 FT TRANSSEM 2393 2415  
 FT TRANSSEM 2416 2438  
 FT TRANSSEM 2439 2461  
 FT TRANSSEM 2462 2484  
 FT TRANSSEM 2485 2507  
 FT TRANSSEM 2508 2530  
 FT TRANSSEM 2531 2553  
 FT TRANSSEM 2554 2576  
 FT TRANSSEM 2577 2599  
 FT TRANSSEM 2600 2622  
 FT TRANSSEM 2623 2645  
 FT TRANSSEM 2646 2668  
 FT TRANSSEM 2669 2691  
 FT TRANSSEM 2692 2714  
 FT TRANSSEM 2715 2737  
 FT TRANSSEM 2738 2760  
 FT TRANSSEM 2761 2783  
 FT TRANSSEM 2784 2806  
 FT TRANSSEM 2807 2829  
 FT TRANSSEM 2830 2852  
 FT TRANSSEM 2853 2875  
 FT TRANSSEM 2876 2898  
 FT TRANSSEM 2899 2921  
 FT TRANSSEM 2922 2944  
 FT TRANSSEM 2945 2967  
 FT TRANSSEM 2968 2990  
 FT TRANSSEM 2991 3013  
 FT TRANSSEM 3014 3036  
 FT TRANSSEM 3037 3059  
 FT TRANSSEM 3060 3082  
 FT TRANSSEM 3083 3105  
 FT TRANSSEM 3106 3128  
 FT TRANSSEM 3129 3151  
 FT TRANSSEM 3152 3174  
 FT TRANSSEM 3175 3197  
 FT TRANSSEM 3198 3220  
 FT TRANSSEM 3221 3243  
 FT TRANSSEM 3244 3266  
 FT TRANSSEM 3267 3289  
 FT TRANSSEM 3290 3312  
 FT TRANSSEM 3313 3335  
 FT TRANSSEM 3336 3358  
 FT TRANSSEM 3359 3381  
 FT TRANSSEM 3382 3404  
 FT TRANSSEM 3405 3427  
 FT TRANSSEM 3428 3450  
 FT TRANSSEM 3451 3473  
 FT TRANSSEM 3474 3496  
 FT TRANSSEM 3497 3519  
 FT TRANSSEM 3520 3542  
 FT TRANSSEM 3543 3565  
 FT TRANSSEM 3566 3588  
 FT TRANSSEM 3589 3611  
 FT TRANSSEM 3612 3634  
 FT TRANSSEM 3635 3657  
 FT TRANSSEM 3658 3680  
 FT TRANSSEM 3681 3703  
 FT TRANSSEM 3704 3726  
 FT TRANSSEM 3727 3749  
 FT TRANSSEM 3750 3772  
 FT TRANSSEM 3773 3795  
 FT TRANSSEM 3796 3817  
 FT TRANSSEM 3818 3840  
 FT TRANSSEM 3841 3863  
 FT TRANSSEM 3864 3886  
 FT TRANSSEM 3887 3909  
 FT TRANSSEM 3910 3932  
 FT TRANSSEM 3933 3955  
 FT TRANSSEM 3956 3978  
 FT TRANSSEM 3979 4001  
 FT TRANSSEM 4002 4024  
 FT TRANSSEM 4025 4047  
 FT TRANSSEM 4048 4070  
 FT TRANSSEM 4071 4093  
 FT TRANSSEM 4094 4115  
 FT TRANSSEM 4116 4137  
 FT TRANSSEM 4138 4160  
 FT TRANSSEM 4161 4183  
 FT TRANSSEM 4184 4206  
 FT TRANSSEM 4207 4229  
 FT TRANSSEM 4230 4252  
 FT TRANSSEM 4253 4275  
 FT TRANSSEM 4276 4298  
 FT TRANSSEM 4299 4321  
 FT TRANSSEM 4322 4344  
 FT TRANSSEM 4345 4367  
 FT TRANSSEM 4368 4390  
 FT TRANSSEM 4391 4413  
 FT TRANSSEM 4414 4436  
 FT TRANSSEM 4437 4459  
 FT TRANSSEM 4460 4482  
 FT TRANSSEM 4483 4505  
 FT TRANSSEM 4506 4528  
 FT TRANSSEM 4529 4551  
 FT TRANSSEM 4552 4574  
 FT TRANSSEM 4575 4597  
 FT TRANSSEM 4598 4620  
 FT TRANSSEM 4621 4643  
 FT TRANSSEM 4644 4666  
 FT TRANSSEM 4667 4689  
 FT TRANSSEM 4690 4712  
 FT TRANSSEM 4713 4735  
 FT TRANSSEM 4736 4758  
 FT TRANSSEM 4759 4781  
 FT TRANSSEM 4782 4804  
 FT TRANSSEM 4805 4827  
 FT TRANSSEM 4828 4850  
 FT TRANSSEM 4851 4873  
 FT TRANSSEM 4874 4896  
 FT TRANSSEM 4897 4919  
 FT TRANSSEM 4920 4942  
 FT TRANSSEM 4943 4965  
 FT TRANSSEM 4966 4988  
 FT TRANSSEM 4989 5011  
 FT TRANSSEM 5012 5034  
 FT TRANSSEM 5035 5057  
 FT TRANSSEM 5058 5080  
 FT TRANSSEM 5081 5103  
 FT TRANSSEM 5104 5126  
 FT TRANSSEM 5127 5149  
 FT TRANSSEM 5150 5172  
 FT TRANSSEM 5173 5195  
 FT TRANSSEM 5196 5218  
 FT TRANSSEM 5219 5241  
 FT TRANSSEM 5242 5264  
 FT TRANSSEM 5265 5287  
 FT TRANSSEM 5288 5310  
 FT TRANSSEM 5311 5333  
 FT TRANSSEM 5334 5356  
 FT TRANSSEM 5357 5379  
 FT TRANSSEM 5380 5402  
 FT TRANSSEM 5403 5425  
 FT TRANSSEM 5426 5448  
 FT TRANSSEM 5449 5471  
 FT TRANSSEM 5472 5494  
 FT TRANSSEM 5495 5517  
 FT TRANSSEM 5518 5540  
 FT TRANSSEM 5541 5563  
 FT TRANSSEM 5564 5586  
 FT TRANSSEM 5587 5609  
 FT TRANSSEM 5610 5632  
 FT TRANSSEM 5633 5655  
 FT TRANSSEM 5656 5678  
 FT TRANSSEM 5679 5701  
 FT TRANSSEM 5702 5724  
 FT TRANSSEM 5725 5747  
 FT TRANSSEM 5748 5770  
 FT TRANSSEM 5771 5793  
 FT TRANSSEM 5794 5816  
 FT TRANSSEM 5817 5839  
 FT TRANSSEM 5840 5862  
 FT TRANSSEM 5863 5885  
 FT TRANSSEM 5886 5908  
 FT TRANSSEM 5909 5931  
 FT TRANSSEM 5932 5954  
 FT TRANSSEM 5955 5977  
 FT TRANSSEM 5978 6000  
 FT TRANSSEM 6001 6023  
 FT TRANSSEM 6024 6046  
 FT TRANSSEM 6047 6069  
 FT TRANSSEM 6070 6092  
 FT TRANSSEM 6093 6115  
 FT TRANSSEM 6116 6138  
 FT TRANSSEM 6139 6161  
 FT TRANSSEM 6162 6184  
 FT TRANSSEM 6185 6207  
 FT TRANSSEM 6208 6230  
 FT TRANSSEM 6231 6253  
 FT TRANSSEM 6254 6276  
 FT TRANSSEM 6277 6299  
 FT TRANSSEM 6300 6322  
 FT TRANSSEM 6323 6345  
 FT TRANSSEM 6346 6368  
 FT TRANSSEM 6369 6391  
 FT TRANSSEM 6392 6414  
 FT TRANSSEM 6415 6437  
 FT TRANSSEM 6438 6460  
 FT TRANSSEM 6461 6483  
 FT TRANSSEM 6484 6506  
 FT TRANSSEM 6507 6529  
 FT TRANSSEM 6530 6552  
 FT TRANSSEM 6553 6575  
 FT TRANSSEM 6576 6598  
 FT TRANSSEM 6599 6621  
 FT TRANSSEM 6622 6644  
 FT TRANSSEM 6645 6667  
 FT TRANSSEM 6668 6690  
 FT TRANSSEM 6691 6713  
 FT TRANSSEM 6714 6736  
 FT TRANSSEM 6737 6759  
 FT TRANSSEM 6760 6782  
 FT TRANSSEM 6783 6805  
 FT TRANSSEM 6806 6828  
 FT TRANSSEM 6829 6851  
 FT TRANSSEM 6852 6874  
 FT TRANSSEM 6875 6897  
 FT TRANSSEM 6898 6920  
 FT TRANSSEM 6921 6943  
 FT TRANSSEM 6944 6966  
 FT TRANSSEM 6967 6989  
 FT TRANSSEM 6990 7012  
 FT TRANSSEM 7013 7035  
 FT TRANSSEM 7036 7058  
 FT TRANSSEM 7059 7081  
 FT TRANSSEM 7082 7104  
 FT TRANSSEM 7105 7127  
 FT TRANSSEM 7128 7150  
 FT TRANSSEM 7151 7172  
 FT TRANSSEM 7173 7195  
 FT TRANSSEM 7196 7218  
 FT TRANSSEM 7219 7241  
 FT TRANSSEM 7242 7264  
 FT TRANSSEM 7265 7287  
 FT TRANSSEM 7288 7310  
 FT TRANSSEM 7311 7333  
 FT TRANSSEM 7334 7356  
 FT TRANSSEM 7357 7379  
 FT TRANSSEM 7380 7402  
 FT TRANSSEM 7403 7425  
 FT TRANSSEM 7426 7448  
 FT TRANSSEM 7449 7471  
 FT TRANSSEM 7472 7494  
 FT TRANSSEM 7495 7517  
 FT TRANSSEM 7518 7540  
 FT TRANSSEM 7541 7563  
 FT TRANSSEM 7564 7586  
 FT TRANSSEM 7587 7609  
 FT TRANSSEM 7610 7632  
 FT TRANSSEM 7633 7655  
 FT TRANSSEM 7656 7678  
 FT TRANSSEM 7679 7701  
 FT TRANSSEM 7702 7724  
 FT TRANSSEM 7725 7747  
 FT TRANSSEM 7748 7770  
 FT TRANSSEM 7771 7793  
 FT TRANSSEM 7794 7816  
 FT TRANSSEM 7817 7839  
 FT TRANSSEM 7840 7862  
 FT TRANSSEM 7863 7885  
 FT TRANSSEM 7886 7908  
 FT TRANSSEM 7909 7931  
 FT TRANSSEM 7932 7954  
 FT TRANSSEM 7955 7977  
 FT TRANSSEM 7978 8000  
 FT TRANSSEM 8001 8023  
 FT TRANSSEM 8024 8046  
 FT TRANSSEM 8047 8069  
 FT TRANSSEM 8070 8092  
 FT TRANSSEM 8093 8115  
 FT TRANSSEM 8116 8138  
 FT TRANSSEM 8139 8161  
 FT TRANSSEM 8162 8184  
 FT TRANSSEM 8185 8207  
 FT TRANSSEM 8208 8230  
 FT TRANSSEM 8231 8253  
 FT TRANSSEM 8254 8276  
 FT TRANSSEM 8277 8299  
 FT TRANSSEM 8300 8322  
 FT TRANSSEM 8323 8345  
 FT TRANSSEM 8346 8368  
 FT TRANSSEM 8369 8391  
 FT TRANSSEM 8392 8414  
 FT TRANSSEM 8415 8437  
 FT TRANSSEM 8438 8460  
 FT TRANSSEM 8461 8483  
 FT TRANSSEM 8484 8506  
 FT TRANSSEM 8507 8529  
 FT TRANSSEM 8530 8552  
 FT TRANSSEM 8553 8575  
 FT TRANSSEM 8576 8598  
 FT TRANSSEM 8599 8621  
 FT TRANSSEM 8622 8644  
 FT TRANSSEM 8645 8667  
 FT TRANSSEM 8668 8690  
 FT TRANSSEM 8691 8713  
 FT TRANSSEM 8714 8736  
 FT TRANSSEM 8737 8759  
 FT TRANSSEM 8760 8782  
 FT TRANSSEM 8783 8805  
 FT TRANSSEM 8806 8828  
 FT TRANSSEM 8829 8851  
 FT TRANSSEM 8852 8874  
 FT TRANSSEM 8875 8897  
 FT TRANSSEM 8898 8920  
 FT TRANSSEM 8921 8943  
 FT TRANSSEM 8944 8966  
 FT TRANSSEM 8967 8989  
 FT TRANSSEM 8990 9012  
 FT TRANSSEM 9013 9035  
 FT TRANSSEM 9036 9058  
 FT TRANSSEM 9059 9081  
 FT TRANSSEM 9082 9104  
 FT TRANSSEM 9105 9127  
 FT TRANSSEM 9128 9150  
 FT TRANSSEM 9151 9172  
 FT TRANSSEM 9173 9195  
 FT TRANSSEM 9196 9218  
 FT TRANSSEM 9219 9241  
 FT TRANSSEM 9242 9264  
 FT TRANSSEM 9265 9287  
 FT TRANSSEM 9288 9310  
 FT TRANSSEM 9311 9333  
 FT TRANSSEM 9334 9356  
 FT TRANSSEM 9357 9379  
 FT TRANSSEM 9380 9402  
 FT TRANSSEM 9403 9425  
 FT TRANSSEM 9426 9448  
 FT TRANSSEM 9449 9471  
 FT TRANSSEM 9472 9494  
 FT TRANSSEM 9495 9517  
 FT TRANSSEM 9518 9540  
 FT TRANSSEM 9541 9563  
 FT TRANSSEM 9564 9586  
 FT TRANSSEM 9587 9609  
 FT TRANSSEM 9610 9632  
 FT TRANSSEM 9633 9655  
 FT TRANSSEM 9656 9678  
 FT TRANSSEM 9679 9701  
 FT TRANSSEM 9702 9724  
 FT TRANSSEM 9725 9747  
 FT TRANSSEM 9748 9770  
 FT TRANSSEM 9771 9793  
 FT TRANSSEM 9794 9816  
 FT TRANSSEM 9817 9839  
 FT TRANSSEM 9840 9862  
 FT TRANSSEM 9863 9885  
 FT TRANSSEM 9886 9908  
 FT TRANSSEM 9909 9931  
 FT TRANSSEM 9932 9954  
 FT TRANSSEM 9955 9977  
 FT TRANSSEM 9978 10000  
 FT TRANSSEM 10001 10023  
 FT TRANSSEM 10024 10046  
 FT TRANSSEM 10047 10069  
 FT TRANSSEM 10070 10092  
 FT TRANSSEM 10093 10115  
 FT TRANSSEM 10116 10138  
 FT TRANSSEM 10139 10161  
 FT TRANSSEM 10162 10184  
 FT TRANSSEM 10185 10207  
 FT TRANSSEM 10208 10230  
 FT TRANSSEM 10231 10253  
 FT TRANSSEM 10254 10276  
 FT TRANSSEM 10277 10299  
 FT TRANSSEM 10300 10322  
 FT TRANSSEM 10323 10345  
 FT TRANSSEM 10346 10368  
 FT TRANSSEM 10369 10391  
 FT TRANSSEM 10392 10414  
 FT TRANSSEM 10415 10437  
 FT TRANSSEM 10438 10460  
 FT TRANSSEM 10461 10483  
 FT TRANSSEM 10484 10506  
 FT TRANSSEM 10507 10529  
 FT TRANSSEM 10530 10552  
 FT TRANSSEM 10553 10575  
 FT TRANSSEM 10576 10598  
 FT TRANSSEM 10599 10621  
 FT TRANSSEM 10622 10644  
 FT TRANSSEM 10645 10667  
 FT TRANSSEM 10668 10690  
 FT TRANSSEM 10691 10713  
 FT TRANSSEM 10714 10736  
 FT TRANSSEM 10737 10759  
 FT TRANSSEM 10760 10782  
 FT TRANSSEM 10783 10805  
 FT TRANSSEM 10806 10828  
 FT TRANSSEM 10829 10851  
 FT TRANSSEM 10852 10874  
 FT TRANSSEM 10875 10897  
 FT TRANSSEM 10898 10920  
 FT TRANSSEM 10921 10943  
 FT TRANSSEM 10944 10966  
 FT TRANSSEM 10967 10989  
 FT TRANSSEM 10990 11012  
 FT TRANSSEM 11013 11035  
 FT TRANSSEM 11036 11058  
 FT TRANSSEM 11059 11081  
 FT TRANSSEM 11082 11104  
 FT TRANSSEM 11105 11127  
 FT TRANSSEM 11128 11150  
 FT TRANSSEM 11151 11172  
 FT TRANSSEM 11173 11195  
 FT TRANSSEM 11196 11218  
 FT TRANSSEM 11219 11241  
 FT TRANSSEM 11242 11264  
 FT TRANSSEM 11265 11287  
 FT TRANSSEM 11288 11310  
 FT TRANSSEM 11311 11333  
 FT TRANSSEM 11334 11356  
 FT TRANSSEM 11357 11379  
 FT TRANSSEM 11380 11402  
 FT TRANSSEM 11403 11425  
 FT TRANSSEM 11426 11448  
 FT TRANSSEM 11449 11471  
 FT TRANSSEM 11472 11494  
 FT TRANSSEM 11495 11517  
 FT TRANSSEM 11518 11540  
 FT TRANSSEM 11541 11563  
 FT TRANSSEM 11564 11586  
 FT TRANSSEM 11587 11609  
 FT TRANSSEM 11610 11632  
 FT TRANSSEM 11633 11655  
 FT TRANSSEM 11656 11678  
 FT TRANSSEM 11679 11701  
 FT TRANSSEM 11702 11724  
 FT TRANSSEM 11725 11747  
 FT TRANSSEM 11748 11770  
 FT TRANSSEM 11771 11793  
 FT TRANSSEM 11794 11816  
 FT TRANSSEM 11817 11839  
 FT TRANSSEM 11840 11862  
 FT TRANSSEM 11863 11885  
 FT TRANSSEM 11886 11908  
 FT TRANSSEM 11909 11931  
 FT TRANSSEM 11932 11954  
 FT TRANSSEM 11955 11977  
 FT TRANSSEM 11978 11999  
 FT TRANSSEM 12000 12023  
 FT TRANSSEM 12024 12046  
 FT TRANSSEM 12047 12069  
 FT TRANSSEM 12070 12092  
 FT TRANSSEM 12093 12115  
 FT TRANSSEM 12116 12138  
 FT TRANSSEM 12139 12161  
 FT TRANSSEM 12162 12184  
 FT TRANSSEM 12185 12207  
 FT TRANSSEM 12208 12230  
 FT TRANSSEM 12231 12253  
 FT TRANSSEM 12254 12276  
 FT TRANSSEM 12277 12299  
 FT TRANSSEM 12300 12322  
 FT TRANSSEM 12323 12345  
 FT TRANSSEM 12346 12368  
 FT TRANSSEM 12369 12391  
 FT TRANSSEM 12392 12414  
 FT TRANSSEM 12415 12437  
 FT TRANSSEM 12438 12460  
 FT TRANSSEM 12461 12483  
 FT TRANSSEM 12484 12506  
 FT TRANSSEM 12507 12529  
 FT TRANSSEM 12530 12552  
 FT TRANSSEM 12553 12575  
 FT TRANSSEM 12576 12598  
 FT TRANSSEM 12599 12621  
 FT TRANSSEM 12622 12644  
 FT TRANSSEM 12645 12667  
 FT TRANSSEM 12668 12690  
 FT TRANSSEM 12691 12713  
 FT TRANSSEM 12714 12736  
 FT TRANSSEM 12737 12759  
 FT TRANSSEM 12760 12782  
 FT TRANSSEM 12783 12805  
 FT TRANSSEM 12806 12828  
 FT TRANSSEM 12829 12851  
 FT TRANSSEM 12852 12874  
 FT TRANSSEM 12875 12897  
 FT TRANSSEM 12898 12920  
 FT TRANSSEM 12921 12943  
 FT TRANSSEM 12944 12966  
 FT TRANSSEM 12967 12989  
 FT TRANSSEM 12990 13012  
 FT TRANSSEM 13013 13035  
 FT TRANSSEM 13036 13058  
 FT TRANSSEM 13059 13081  
 FT TRANSSEM 13082 13104  
 FT TRANSSEM 13105 13127  
 FT TRANSSEM 13128 13150  
 FT TRANSSEM 13151 13172  
 FT TRANSSEM 13173 13195  
 FT TRANSSEM

Query Match 57.9%; Score 33; DB 1; Length 462;  
 Best Local Similarity 62.5%; Pred. No. 27;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 PAPMKFF 9  
 | | | | |  
 Db 415 PGPDKFF 422

## RESULT 15

PT1\_YEAST STANDARD; PRT: 476 AA.  
 AC P09368;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Proline oxidase, mitochondrial precursor (EC 1.5.3.-) (Proline dehydrogenase).  
 GN PT1 OR YLR142W OR L3170 OR L9606.2.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88142835; PubMed=3125423;  
 RA Wang S.-S., Brandtiss M.C.;  
 RT "Proline utilization in Saccharomyces cerevisiae: sequence, regulation, and mitochondrial localization of the PT1 gene product."  
 RT Mol. Cell. Biol. 7:4431-4440(1987).  
 RL [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C;  
 RA Delius H., Hedling U.;  
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Rieger M., Mueller-Auer S., Brueckner M.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Favello A., Fulton L., Gattung S., Greco T., Kirsten J., Kucaba T., Hallsworth K., Hawkins J., Hillier L., Jier M., Johnson D., Johnston L., Langston Y., Latreille P., Le T., Mardis E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D., Rifken L., Riles L., Raich A., Trevaskis E., Vignati D., Wilcox L., Wohlman P., Vaudin M., Wilson R., Waterston R.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: CONVERTS PROLINE TO DELTA-1-PYRROLINE-5-CARBOXYLATE.  
 CC -1- PATHWAY: FIRST STEP IN THE CONVERSION FROM PROLINE TO GLUTAMATE.  
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.  
 CC -1- MISCELLANEOUS: THE EXPRESSION OF PROLINE OXIDASE IS CONTROLLED BY PROLINE AND OXYGEN.  
 CC -1- MISCELLANEOUS: PROLINE OXIDASE REQUIRES A FUNCTIONAL ELECTRON TRANSPORT CHAIN AEROBICOSIS FOR ITS ACTIVITY.  
 CC -1- SIMILARITY: BELONGS TO THE PROLINE OXIDASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M18107; AAA16631.1; -  
 DR EMBL; X91258; CAA62662.1; -  
 DR EMBL; Z73314; CAA97714.1; -  
 DR EMBL; U53881; AAB82390.1; -  
 DR PIR; A29359; A29359.

DR SGD; S0004132; PT1.  
 DR InterPro; IPR002872; Pro\_dh.  
 DR Pfam; PF01619; Pro\_dh; 1.  
 KW Oxidoreductase; Proline metabolism; Mitochondrion; Transit peptide.  
 FT TRANSIT 1 ?  
 FT CHAIN ?  
 FT CONFLICT 50 476  
 FT CONFLICT 185 185  
 SQ SEQUENCE 476 AA; 53271 MW; 12FE55F86E107D CRC64;

Query Match 57.9%; Score 33; DB 1; Length 476;  
 Best Local Similarity 50.0%; Pred. No. 28;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 PAPMKFFTT 11  
 | | | | |  
 Db 100 PIPVKEFVS 109

Search completed: June 11, 2002, 15:54:49  
 Job time: 203 sec



**This Page Blank (uspto)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2002, 15:50:51 ; Search time 33.9 Seconds  
(without alignments)  
61.237 Million cell updates/sec

Title: US-09-727-198-5  
Perfect score: 57  
Sequence: 1 XPAPMKFFPTX 12

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_19:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mmc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriaph:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	66.7	120	3	013904
2	37	64.9	598	17	0973K4
3	36	63.2	225	5	017989
4	36	63.2	300	10	09SUS5
5	36	63.2	353	5	09XHP5
6	35	63.2	519	10	09AXL7
7	35	61.4	121	10	09SL50
8	35	61.4	163	4	09H7F5
9	35	61.4	166	11	09C018
10	35	61.4	189	17	09V243
11	35	61.4	191	17	057989
12	35	61.4	222	3	008278
13	35	61.4	292	6	09GME2
14	35	61.4	328	10	09SAD4
15	35	61.4	388	5	09X73
16	35	61.4	443	16	050980

17	35	61.4	476	16	09H2G1
18	35	61.4	810	3	012172
19	35	61.4	2016	5	09VLS7
20	35	61.4	6315	2	09ADL6
21	34	59.6	162	10	09SKN0
22	34	59.6	253	5	09SRJ4
23	34	59.6	274	5	09XVQ9
24	34	59.6	284	10	09SIU3
25	34	59.6	336	10	048824
26	34	59.6	373	2	057286
27	34	59.6	486	10	09FM41
28	34	59.6	487	16	067873
29	34	59.6	931	10	09M9A7
30	33	57.9	202	4	09BX48
31	33	57.9	246	5	09S5G2
32	33	57.9	284	10	09LJD3
33	33	57.9	312	10	09FV16
34	33	57.9	325	16	066707
35	33	57.9	362	3	001785
36	33	57.9	371	3	001787
37	33	57.9	371	3	006326
38	33	57.9	455	3	094738
39	33	57.9	462	4	09HB49
40	33	57.9	475	5	09VB80
41	33	57.9	491	5	09VB81
42	33	57.9	640	5	09VF65
43	33	57.9	1331	4	096PE1
44	33	57.9	1967	5	09VPV1
45	32.5	57.0	133	12	084462

## ALIGNMENTS

RESULT 1

ID	013904	PRELIMINARY;	PRT;	120 AA.
AC	013904;			
DT	01-JAN-1999 (TREMBLrel. 09, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	HYPOTHELICAL 13.9 KDA PROTEIN C22A12.13 IN CHROMOSOME I.			
GN	SPAC22A12.13.			
OS	Schizosaccharomyces pombe (Fission yeast).			
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;			
OC	Schizosaccharomycetales; Schizosaccharomycetaceae;			
OX	NCBI_TaxID=4896;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=972;			
RA	Pearson D., Churcher C.M., Wood V., Barrell B.G., Rajandream M.A.;			
RL	Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.			
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).			
DR	EMBL; Z99295; CAB16583.2; -			
KW	Hypothetical protein; Transmembrane.			
FT	TRANSMEM 10 30 POTENTIAL.			
FT	TRANSMEM 48 68 POTENTIAL.			
SQ	SEQUENCE 120 AA; 13903 MW; EB9431495A6004B2 CRC64;			

Query Match 66.7%; Score 38; DB 3; Length 120;  
Best Local Similarity 75.0%; Pred. No. 2.5;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 PAPMKFF 9  
Db 32 PAPVLFKFF 39

RESULT 2  
ID 0973K4 PRELIMINARY; PRT; 598 AA.

AC Q973K4;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
 DE PUTATIVE THERMOPROTEIN.  
 GN ST0894.  
 OS Sulfolobus tokodaii.  
 OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.  
 RX NCBI\_TaxID=111955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JCM 10545 / 7;  
 RX PubMed=11572479;  
 RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,  
 RA Sekine M., Baba S.-I., Anka A., Kosugi H., Hosoyama A., Fukui S.,  
 RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Kato Y.,  
 RA Yoshizawa T., Tanaka T., Otsuka Y., Yamazaki J., Kushida N., Oguchi A.,  
 RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,  
 RA Oshima T., Kikuchi H.;  
 RT "Complete genome sequence of an aerobic thermophilic  
 RT Crenarchaeon, Sulfolobus tokodaii strain7.";  
 RL DNA Res. 8:123-140(2001).  
 DR EMBL: AP000984; BAB5908.1; -  
 KW Hypothetical protein: Complete proteome.  
 SQ SEQUENCE 598 AA; 68102 MW; 20E848698BB931F6 CRC64;

Query Match 64.9%; Score 37; DB 17; Length 598;  
 Best Local Similarity 75.0%; Pred. No. 20;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 PPMKFFTT 11  
 11:||||  
 DB 271 PVMRFFTT 278

RESULT 3  
 ID 017989 PRELIMINARY; PRT; 225 AA.  
 AC 017989;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
 DE R08H2.6 PROTEIN.  
 GN R08H2.6.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Matthews L.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP MEDLINE=99069613; PubMed=9851916;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology.";  
 RL Science 282:2012-2018(1998).  
 DR EMBL: Z81575; CAB04632.1; -  
 DR InterPro: IPR003003; 7TM\_chemorecept\_2.  
 DR InterPro: IPR000168; 7TM\_nematode.  
 DR Pfam: PF01604; 7tm\_5; 1.  
 SQ SEQUENCE 225 AA; 25500 MW; BFF5F5E0D5313D4 CRC64;

Query Match 63.2%; Score 36; DB 5; Length 225;  
 Best Local Similarity 60.0%; Pred. No. 12;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 PPMKFFTT 11  
 11:||||

DB 77 PCPAGEFTT 86  
 RESULT 4  
 ID 09S055 PRELIMINARY; PRT; 300 AA.  
 AC 09S055;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, last annotation update)  
 DE HYPOTHETICAL 34.8 KDA PROTEIN.  
 GN T17F15.210.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Choisne N., Robert C., Brothier P., Wincker P., Catolico L.,  
 RA Artiguenave F., Saurin W., Weissenbach J., Mewes H.W., Mayer K.F.X.,  
 RA Lemcke K., Schueller C., Quetier F., Salanoubat M.;  
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL049658; CAB41147.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 300 AA; 34755 MW; 750C980DF9F94B355 CRC64;

Query Match 63.2%; Score 36; DB 10; Length 300;  
 Best Local Similarity 60.0%; Pred. No. 16;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 PPMKFFTT 11  
 11:||||  
 DB 40 PAVLNEFTT 49

RESULT 5  
 ID 09XXP5 PRELIMINARY; PRT; 353 AA.  
 AC 09XXP5;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
 DE F08E10.3 PROTEIN.  
 GN F08E10.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Gardner A.E.;  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology.";  
 RL Science 282:2012-2018(1998).  
 DR EMBL: AL021566; CA16499.1; -  
 DR InterPro: IPR003003; 7TM\_chemorecept\_2.  
 DR InterPro: IPR000168; 7TM\_nematode.  
 DR Pfam: PF01604; 7tm\_5; 1.  
 SQ SEQUENCE 353 AA; 39782 MW; 87D126B242BF754E CRC64;

Query Match 63.2%; Score 36; DB 5; Length 353;

Best Local Similarity 60.0%; Pred. No. 19;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
OY 2 PAPMKFFTT 11  
11111111  
Db 187 PPTPTPTFT 196

RESULT 6  
O9AXL7 PRELIMINARY; PRT: 519 AA.  
AC O9AXL7; PRT: 519 AA.  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE MALATE SYNTHASE (EC 4.1.3.2) (FRAGMENT).  
OS Musa acuminata (Banana).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Zingiberales; Musaceae;  
OC Musa.  
OX NCBI\_TaxID=4641;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV, WILLIAMS;  
RA Chandramouli S., Pua E.C.;  
RT "Molecular cloning and characterization of malate synthase in banana  
(Musa acuminata).";  
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: L-MALATE + COA -> ACETYL-COA + H(2)O +  
GLYOXALATE.  
CC -1- PATHWAY: SECOND STEP IN GLYOXALATE BYPASS, AN ALTERNATIVE TO THE  
TRICARBOXYLIC ACID CYCLE (IN BACTERIA, FUNGI AND PLANTS).  
CC -1- SIMILARITY: BELONGS TO THE MALATE SYNTHASE FAMILY.  
DR EMBL; AF31286; AK07428.1; -  
DR InterPro; IPR001465; Malate\_synthase.  
DR PROSITE; PS00510; MALATE\_SYNTHASE.1.  
KM Glyoxylate bypass; Lyase; Tricarboxylic acid cycle.  
FT NON\_TER 1  
SQ SEQUENCE 519 AA; 59224 MW; 854A3C4B5768661 CRC64;

Query Match 63.2%; Score 36; DB 10; Length 519;  
Best Local Similarity 75.0%; Pred. No. 28;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
OY 4 PMMKFFTT 11  
11111111  
Db 346 PIMKFTTS 353

RESULT 7  
O9SL50 PRELIMINARY; PRT: 121 AA.  
AC O9SL50;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE AT2G17960 PROTEIN.  
GN AT2G17960.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eustosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV, COLUMBIA;  
RA MEDLINE=20083487; PubMed=10617197;  
RA Lin X., Kaul S., Roumsey S.D., Shea T.P., Benito M.-I., Town C.D.,  
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.B., Feldblyum T.V.,  
RA Bueli C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,  
RA Cronin L.A., Shen M., Vankken S.E., Umayam L., Tallon L.J., Gill J.E.,  
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,

RA Copenhagen G.P., Preuss D., Niernan W.C., White O., Eisen J.A.,  
RA Salzberg S.L., Fraser C.M., Venter J.C.;  
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
thaliana";  
RL Nature 402:761-768(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV, COLUMBIA;  
RA Lin X.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC006201; AAD20130.1; -  
SQ SEQUENCE 121 AA; 14055 MW; F8879F2C984753C0 CRC64;

Query Match 61.4%; Score 35; DB 10; Length 121;  
Best Local Similarity 50.0%; Pred. No. 11;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
OY 2 PAPMKFFTT 11  
11111111  
Db 68 PLPLQOYFTT 77

RESULT 8  
O9H7F5 PRELIMINARY; PRT: 163 AA.  
AC O9H7F5;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE CDNA: FLJ20974 FIS, CLONE ADS001596.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=ADIPOSE TISSUE;  
RA Kawakami T., Nozuchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,  
RA Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,  
RA Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Ota T., Suzuki Y.,  
RA Oda-yashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,  
RA Isogai T., Sugano S.;  
RT "NEBO human cDNA sequencing project.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK024627; BAB14937.1; -  
SQ SEQUENCE 163 AA; 17712 MW; 7CD8D8306DEB1EDC CRC64;

Query Match 61.4%; Score 35; DB 4; Length 163;  
Best Local Similarity 66.7%; Pred. No. 14;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
OY 2 PAPMKFFTT 10  
11111111  
Db 41 PAVVGRFTT 49

RESULT 9  
O9CQ18 PRELIMINARY; PRT: 166 AA.  
AC O9CQ18;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE 1500026D16RIK PROTEIN.  
GN 1500026D16RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN-C57BL/6J; TISSUE-CEREBELLUM, AND EMBRYO;  
 RX MEDLINE-21085660; PubMed-11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaishi I., Pesole G., Quackenbush J.,  
 RA Schiml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Brownstein M., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Norioka P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmung L.,  
 RA Wyszynski B., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,  
 RA Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK018776; BAB31401.1; -;  
 DR EMBL: AK011556; BAB27694.1; -;  
 DR MGD: MGI:1915459; 1500026D16Rik.  
 SQ SEQUENCE 166 AA; 17823 MW; D631CFE252040417 CRC64;

Query Match 61.4%; Score 35; DB 11; Length 166;  
 Best Local Similarity 66.7%; Pred. No. 15;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 PAPMKFFT 10  
 Db 41 PAPVDRFT 49

RESULT 10  
 Q9V243 PRELIMINARY; PRT; 189 AA.  
 AC Q9V243;  
 DT 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)  
 DE HYPOTHETICAL 20.9 KDA PROTEIN.  
 GN PAB2198.  
 OS Pyrococcus abyssi.  
 OC Archaea: Euryarchaeota; Thermococcales; Pyrococcus.  
 OX NCBI\_TaxID=29292;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ORSAY;  
 RA Helling R.;  
 RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure and evolution."  
 RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ248283; CAB49155.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 189 AA; 20866 MW; 47177FID072EA5B CRC64;

Query Match 61.4%; Score 35; DB 17; Length 189;  
 Best Local Similarity 75.0%; Pred. No. 17;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 PMMKFFT 11  
 Db 78 PYMKFFAT 85

RESULT 11  
 Q57989 PRELIMINARY; PRT; 191 AA.

AC 057989;  
 DT 01-AUG-1998 (TREMblrel. 07, Created)  
 DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)  
 DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)  
 DE HYPOTHETICAL 21.2 KDA PROTEIN PH0251.  
 GN PH0251.  
 OS Pyrococcus horikoshii.  
 OC Archaea: Euryarchaeota; Thermococcales; Pyrococcus.  
 OX NCBI\_TaxID=53953;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-OT3;  
 RX MEDLINE-98344137; PubMed-9679194;  
 RA Kawarabayashi Y., Sawada M., Horikawa H., Halkawa Y., Hino Y.,  
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,  
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,  
 RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,  
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,  
 RA Masuchi Y., Shizuya H., Kikuchi H.;  
 RT "Complete sequence and gene organization of the genome of a hyper-thermophilic archaeobacterium, Pyrococcus horikoshii OT3."  
 RL DNA Res. 5:55-76(1998).  
 DR EMBL: AP000001; BAA29323.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 191 AA; 21150 MW; F7CE582B4F6A0BDE CRC64;

Query Match 61.4%; Score 35; DB 17; Length 191;  
 Best Local Similarity 75.0%; Pred. No. 17;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 PMMKFFT 11  
 Db 78 PYMKFFAT 85

RESULT 12  
 Q08278 PRELIMINARY; PRT; 222 AA.  
 AC Q08278;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE CHROMOSOME XV READING FRAME ORF YOL135C.  
 GN MED7 OR YOL135C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Arino J., Casamayor A., Gamo F.J., Gancedo C., Lafuente M.J.,  
 RA Aldea M., Casas C., Herrero E.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA WPS;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-FY1679;  
 RX MEDLINE-97051593; PubMed-8896270;  
 RA Aldea M., Piedrafita L., Casas C., Casamayor A., Khalid H.,  
 RA Balcells L., Arino J., Herrero E.;  
 RT "Sequence analysis of a 12 801 bp fragment of the left arm of yeast chromosome XV containing a putative 6-phosphofructo-2-kinase gene, a gene for a possible glycopospholipid-anchored surface protein and six other open reading frames."  
 RT Yeast 12:1053-1058(1996).  
 RL EMBL: Z74877; CAA99156.1; -;  
 DR EMBL: X95465; CAA64734.1; -;  
 DR SGD: S0005495; MED7.  
 SQ SEQUENCE 222 AA; 25585 MW; A883014EF3223C9 CRC64;



```

Query Match                      61.4%; Score 35; DB 3; Length 222;
Best Local Similarity 66.7%; Pred. No. 20;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PAPMKFFTT 10
   1 : : : : :
Db 16 PPVYKFFTT 24

RESULT 13
O9GME2 PRELIMINARY; PRT; 292 AA.
AC O9GME2:
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DE 17-BETA-HYDROXYSTEROID DEHYDROGENASE TYPE 1 (FRAGMENT).
OS Callithrix jacchus (Common marmoset).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Callithrix.
OX NCBI_TaxID=9483;

RN [1]
RP SEQUENCE FROM N.A.
RA Hussen B., Einspanier A.;
RT "Regulation of 17beta-hydroxysteroid dehydrogenases type 1 and type 7
   during pregnancy in the marmoset monkey.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
   (SDR) FAMILY.
DR EMBL: AF272013; AAC01115.2; -.
DR HSSP: P14061; IEDS.
DR InterPro: IPR002198; ADH_short.
DR Pfam: PF00106; adh_short.1.
DR PRINTS: PR00080; SDRFAMILY.
DR PROSITE: PS00061; ADH_SHORT; UNKNOWN_1.
KW Oxidoreductase.
FT NON_TER 1
SQ SEQUENCE 292 AA; 31577 MW; A815EE41A51B43FC CRC64;

Query Match                      61.4%; Score 35; DB 6; Length 292;
Best Local Similarity 50.0%; Pred. No. 26;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PAPMKFFTT 11
   1 : : : : :
Db 226 PKPALRYFTT 235

RESULT 14
O9SAD4 PRELIMINARY; PRT; 328 AA.
AC O9SAD4:
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE F3rl9.1 PROTEIN.
GN F3rl9.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;

RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CV, COLUMBIA;
RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lee J.M., Lenz C.,
   Liu S., Li J., Kremenetskaia I., Luros J., Ngan I., Gonzalez A.,
   Altafi H., Araujo R., Chao Q., Conn L., Conway A.B., Dunn P.,
   Hansen N., Huizer L., Kim C., Palm C., Rowley D., Shinn P., Walker M.,
   Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.;

```

```

RT "Arabidopsis thaliana chromosome 1 BAC F3rl9 sequence.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=CV, COLUMBIA;
RC Theologis A.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC007357; AAD31052.1; -.
DR HSSP: O80337; 2GCC.
DR InterPro: IPR001471; AP2-domain.
DR Pfam: PF00847; AP2-domain; 1.
DR PRINTS: PR00367; ETHRSPLENT.
DR Prodom: PD001423; AP2-domain; 1.
DR SMART: SM00380; AP2; 1.
SQ SEQUENCE 328 AA; 36319 MW; 66320534B5FCAE6B CRC64;

Query Match                      61.4%; Score 35; DB 10; Length 328;
Best Local Similarity 66.7%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PAPMKFFTT 10
   1 : : : : :
Db 17 PNPTRKFFTT 25

RESULT 15
O9XX73 PRELIMINARY; PRT; 388 AA.
AC O9XX73:
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Y102A5C.21 PROTEIN.
GN Y102A5C.21.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;

RN [1]
RP SEQUENCE FROM N.A.
RA Gardner A.E.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
   investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL: AL031627; CAA20962.1; -.
DR InterPro: IPR003003; 7TM_chemorecept_2.
DR InterPro: IPR00168; 7TM_nematode.
DR Pfam: PF01604; 7tm_5; 1.
SQ SEQUENCE 388 AA; 44245 MW; B50BB2CFE2F9B77E CRC64;

Query Match                      61.4%; Score 35; DB 5; Length 388;
Best Local Similarity 60.0%; Pred. No. 34;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 PAPMKFFTT 11
   1 : : : : :
Db 220 PCPTRFFFTT 229

```

Search completed: June 11, 2002, 15:54:24  
Job time: 213 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2002, 15:47:49 ; Search time 15.04 Seconds  
(without alignments)  
76.667 Million cell updates/sec

Title: US-09-727-198-5

Perfect score: 57

Sequence: 1 XPAPMMKFFTX 12

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 segs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.1:  
2: PIR.2:  
3: PIR.3:  
4: PIR.4:

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	66.7	157	2 T38153	hypothetical prote
2	37	64.9	285	2 AD1992	hypothetical prote
3	36	63.2	225	2 T24052	hypothetical prote
4	36	63.2	300	2 T06691	hypothetical prote
5	36	63.2	353	2 T20580	hypothetical prote
6	36	63.2	485	2 AD2372	hypothetical prote
7	35	61.4	121	2 E84558	hypothetical prote
8	35	61.4	189	2 D75213	hypothetical prote
9	35	61.4	191	2 D71249	hypothetical prote
10	35	61.4	222	2 S66832	hypothetical prote
11	35	61.4	304	2 T16704	hypothetical prote
12	35	61.4	328	1 DEH07	hypothetical prote
13	35	61.4	328	2 G86263	estradiol 17beta-d
14	35	61.4	388	2 T26360	hypothetical prote
15	35	61.4	443	2 H70216	hypothetical prote
16	35	61.4	476	2 H83265	PTS system, cellob
17	35	61.4	810	2 S67050	probable D-alanyl-
18	34	59.6	162	2 D84684	hypothetical prote
19	34	59.6	274	2 T20786	hypothetical prote
20	34	59.6	278	2 A49067	transcription init
21	34	59.6	284	2 B84508	probable Tail-like
22	34	59.6	336	2 T00832	probable Tail-like
23	34	59.6	373	2 E49094	methylinonyl-CoA
24	34	59.6	453	2 S18597	tubulin beta chain
25	34	59.6	487	2 E70480	probable sodium/pr
26	34	59.6	931	2 H96527	protein F27J15.16
27	33	57.9	180	2 PS0011	lignin peroxidase
28	33	57.9	180	2 PS0011	lignin peroxidase
29	33	57.9	273	1 B27035	acid phosphatase (

30	33	57.9	325	2 G70334	hypothetical prote
31	33	57.9	371	1 OPUG65	diarylopropane pero
32	33	57.9	371	2 JCI1268	lignin peroxidase
33	33	57.9	372	1 OPUGAP	lignin peroxidase
34	33	57.9	372	1 OPUGBP	lignin peroxidase
35	33	57.9	372	1 OPUG3P	lignin peroxidase
36	33	57.9	372	2 B32322	lignin peroxidase
37	33	57.9	372	2 A36693	lignin peroxidase
38	33	57.9	372	2 A43638	lignin peroxidase
39	33	57.9	372	2 T04002	lignin peroxidase
40	33	57.9	372	2 A32322	lignin peroxidase
41	33	57.9	372	2 S01028	lignin peroxidase
42	33	57.9	372	2 JH0156	lignin peroxidase
43	33	57.9	372	2 S69246	lignin peroxidase
44	33	57.9	373	2 A27817	lignin peroxidase
45	33	57.9	461	2 A31237	alpha-2C-adrenergic

#### ALIGNMENTS

RESULT 1  
T38153  
hypothetical protein SPAC22A12.13 - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T38153  
R:Pearson, D.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
Submitted to the EMBL Data Library, September 1997  
A:Reference number: 221774  
A:Accession: T38153  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-157 <PEAK>  
A:Cross-references: EMBL:Z99295; PIDN:CAB15683.1; GSPDB:GN00066; SPDB:SPAC22A12.13  
A:Experimental source: Strain 972h-; cosmid c22A12  
C:Genetics:  
A:Gene: SPDB:SPAC22A12.13  
A:Map position: 1

Query Match 66.7%; Score 38; DB 2; Length 157;  
Best Local Similarity 75.0%; Pred. No. 2.2;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 2 PAPMMKFF 9.  
DB 69 PAPVLRFF 76

RESULT 2  
AD1992  
hypothetical protein all1489 [imported] - Anabaena sp. (strain PCC 7120)  
C:Species: Anabaena sp.  
A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 11-Jan-2002  
R:Keneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AD1992  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-285 <KOR>  
A:Cross-references: GB:BA000019; PIDN:BA077855.1; PID:q17135310; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: all1489

Query Match 64.9%; Score 37; DB 2; Length 285;

Best Local Similarity 60.0%; Pred. No. 6.7;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 PAPMKFFTT 11  
|||: |||  
Db 60 PAPVORFFRT 69

## RESULT 3

T24052

hypothetical protein R08H2.6 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T24052

R:Matthews, L.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19835

A:Accession: T24052

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-225 <MIL>

A:Cross-References: EMBL:Z81575; PIDN:CA804632.1; GSPDB:GN00023; CESP:R08H2.6

A:Experimental source: clone R08H2

C:Genetics:

A:Gene: CESP:R08H2.6

A:Map position: 5

Query Match 63.2%; Score 36; DB 2; Length 225;

Best Local Similarity 60.0%; Pred. No. 8.4;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 PAPMKFFTT 11  
| | : ||||  
Db 77 PCPADFFTT 86

## RESULT 4

T06691

hypothetical protein T17F15.210 - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 24-May-2001

C:Accession: T06691

R:Quetier, F.; Choisme, N.; Robert, C.; Brotier, P.; Wincker, P.; Catolico, L.; Artigou

submitted to the Protein Sequence Database, April 1999

A:Reference number: Z15793

A:Accession: T06691

A:Molecule type: DNA

A:Residues: 1-300 <QUB>

A:Cross-References: EMBL:AL049658; GSPDB:GN00061; ATSP:T17F15.210

A:Experimental source: cultivar Columbia; BAC clone T17F15

C:Genetics:

A:Gene: ATSP:T17F15.210

A:Map position: 3

C:Superfamily: *Arabidopsis thaliana* hypothetical protein At2g17920

Query Match 63.2%; Score 36; DB 2; Length 300;

Best Local Similarity 60.0%; Pred. No. 11;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 PAPMKFFTT 11  
|||: |||  
Db 40 PAPVLFVTT 49

## RESULT 5

T20580

hypothetical protein F08E10.3 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000

C:Accession: T20580

R:Gardner, A.

submitted to the EMBL Data Library, January 1998

A:Reference number: Z19296

A:Accession: T20580

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-353 <MIL>

A:Cross-References: EMBL:AL021566; PIDN:CAA16499.1; GSPDB:GN00023; CESP:F08E10.3

A:Experimental source: clone F08E10

C:Genetics:

A:Gene: CESP:F08E10.3

A:Map position: 5

A:Introns: 121/1; 332/1

C:Superfamily: *Caenorhabditis* hypothetical protein C4967.2

Query Match 63.2%; Score 36; DB 2; Length 353;

Best Local Similarity 60.0%; Pred. No. 13;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 PAPMKFFTT 11  
| | : ||||  
Db 187 PCPTORFFTT 196

## RESULT 6

AD2372

hypothetical protein alr4532 [imported] - *Anabaena* sp. (strain PCC 7120)

C:Species: *Anabaena* sp.

A:Note: *Anabaena* sp. (strain PCC 7120) is a synonym of *Nostoc* sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 11-Jan-2002

C:Accession: AD2372

R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kurita, T.; Sasamoto, S.; Watanabe, A.; Iriugu

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AD2372

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-485 <KUR>

A:Cross-References: GB:BA000019; PIDN:BA876231.1; PID:q17133668; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: alr4532

Query Match 63.2%; Score 36; DB 2; Length 485;

Best Local Similarity 75.0%; Pred. No. 19;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 APMKFFTT 10  
|||: |||  
Db 207 APLMEFFTT 214

## RESULT 7

E84558

hypothetical protein At2g17960 [imported] - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C:Accession: E84558

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,

neuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A:Reference number: AB4420; MUID:20083487

A:Accession: E84558

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-121 <STO>

A:Cross-References: GB:AE002093; NID:g44406822; PIDN:AA020130.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g17960  
A:Map position: 2

Query Match  
Best Local Similarity 61.4%; Score 35; DB 2; Length 121;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 PAMMKFFTT 11  
| : : : |||  
Db 68 PLPLQYFTT 77

RESULT 8

D75213  
hypothetical protein PAB2198 - Pyrococcus abyssi (strain Orsay)

C:Species: Pyrococcus abyssi

C>Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Aug-1999

C:Accession: D75213

R:Anonymous, Genoscope

Submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru

A:Reference number: A75001

A:Accession: D75213

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-189 <RAW>

A:Cross-references: GB:AJ248283; GB:AL096836; NID:g5457433; PIDN:CAM49155.1; PID:el51504

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: PAB2198

Query Match  
Best Local Similarity 61.4%; Score 35; DB 2; Length 189;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PAMMKFFTT 11  
| : : : |||  
Db 78 PVMKFFAT 85

RESULT 9

D71249  
hypothetical protein PH0251 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii

C>Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 21-Jul-2000

C:Accession: D71249

R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Seki

M.; Ohnuki, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kishida, N.; Oguchi

DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic

A:Reference number: A71000; MUID:98344137

A:Accession: D71249

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-191 <RAW>

A:Cross-references: GB:AP000001; NID:g3236128; PIDN:BA29323.1; PID:g3256640

A:Experimental source: strain OT3

A:Note: this accession replaces an interim accession for a sequence replaced by GenBank

C:Genetics:

A:Gene: PH0251

Query Match  
Best Local Similarity 61.4%; Score 35; DB 2; Length 191;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PAMMKFFTT 11  
| : : : |||  
Db 78 PVMKFFAT 85

RESULT 10

S66832  
hypothetical protein YOL135c - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein AOE222; hypothetical protein 00505

C:Species: Saccharomyces cerevisiae

C>Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 21-Jul-2000

C:Accession: S66832; S72032

R:Ariño, J.; Casamayor, A.; Gamio, F.J.; Gancedo, C.; Lafuente, M.J.; Aldea, M.; Casas

submitted to the Protein Sequence Database, July 1996

A:Reference number: S66814

A:Accession: S66832

A:Molecule type: DNA

A:Residues: 1-222 <ART>

A:Cross-references: EMBL:Z74877; NID:g1420025; PID:e252305; PID:g1420026; MIPS:YOL135

A:Experimental source: strain S288C

R:Aldea, M.; Piedrafito, L.; Casas, C.; Casamayor, A.; Khalid, H.; Balcells, L.; Arin

yeast 12, 1053-1058, 1996

A:Title: Sequence analysis of a 12 801 bp fragment of the left arm of yeast chromosome

protein and six other open reading frames.

A:Reference number: S72030; MUID:97051593

A:Accession: S72032

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-222 <ALD>

A:Cross-references: EMBL:X95465; NID:g1628437; PIDN:CAA64734.1; PID:g1628440

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1996

C:Genetics:

A:Map position: 15L

A:Note: YOL135c

Query Match  
Best Local Similarity 66.7%; Score 35; DB 2; Length 222;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PAMMKFFTT 10  
| : : : |||  
Db 16 PPPYKFFTT 24

RESULT 11

T16704  
hypothetical protein R07G3.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C:Accession: T16704

R:Connell, M.

submitted to the EMBL Data Library, July 1995

A:Description: The sequence of C. elegans cosmid R07G3.

A:Reference number: Z18562

A:Accession: T16704

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-304 <CON>

A:Cross-references: EMBL:U23452; NID:g733564; PID:g733569; PIDN:AAC46751.1; CESP:R07G

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:R07G3.6

A:introns: 26/3

Query Match  
Best Local Similarity 75.0%; Score 35; DB 2; Length 304;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PAMMKFFTT 11  
| : : : |||  
Db 235 PLMKFFTT 242

RESULT 12

DEHUE7

estradiol 17beta-dehydrogenase (EC 1.1.1.62) type 1 [validated] - human

N:Alternate names: 17beta-estradiol dehydrogenase; 17beta-hydroxysteroid dehydrogenase  
 C:Species: Homo sapiens (man)  
 C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 08-Dec-2000  
 C:Accession: A36081; S29288; S01654; S43212; A40146; A60440; A21031; A60628; G02067; A37  
 R:Luu-The, V.; Labrie, C.; Sismard, J.; Lachance, Y.; Zhao, H.F.; Couet, J.; Leblanc, G.;  
 Mol. Endocrinol. 4, 268-275, 1990  
 A:Title: Structure of two in tandem human 17beta-hydroxysteroid dehydrogenase genes.  
 A:Reference number: A36081; MUID:90231340  
 A:Accession: A36081  
 A:Molecule type: DNA  
 A:Residues: 1-328 <LNU>  
 A:Cross-references: GB:M27138; NID:9181950; PIDN:AAB16941.1; PID:9181951  
 R:Beloketo, H.; Isomaa, V.; Viikho, R.  
 Eur. J. Biochem. 209, 459-466, 1992  
 A:Title: Genomic organization and DNA sequences of human 17beta-hydroxysteroid dehydroge  
 A:Reference number: S29288; MUID:93011163  
 A:Accession: S29288  
 A:Molecule type: DNA  
 A:Residues: 1-328 <PEL>  
 A:Cross-references: EMBL:M84472; NID:9806392; PIDN:AAB16942.1; PID:9177127  
 A:Note: the authors did not translate the codon for residue 1  
 A:Note: the authors translated the codon AAC for residue 233 as Asp  
 R:Beloketo, H.; Isomaa, V.; Maentausta, O.; Viikho, R.  
 FEBS Lett. 239, 73-77, 1988  
 A:Title: Complete amino acid sequence of human placental 17-beta-hydroxysteroid dehydrog  
 A:Reference number: S01654; MUID:89031223  
 A:Accession: S01654  
 A:Molecule type: mRNA  
 A:Residues: 2-328 <PEL2>  
 A:Cross-references: GB:X13440; NID:923364; PIDN:CAA31792.1; PID:923365  
 A:Accession: S43212  
 A:Molecule type: protein  
 A:Residues: 2-24; 53-68; 205-223 <PE2>  
 R:Lu, The, V.; Labrie, C.; Zhao, H.F.; Couet, J.; Lachance, Y.; Sismard, J.; Leblanc, G.;  
 Mol. Endocrinol. 3, 1301-1309, 1989  
 A:Title: Characterization of cDNAs for human estradiol 17beta-dehydrogenase and assignme  
 A:Reference number: A40146; MUID:89384667  
 A:Accession: A40146  
 A:Molecule type: mRNA  
 A:Residues: 1-328 <THE>  
 A:Cross-references: GB:M27138; NID:9181950; PIDN:AAB16941.1; PID:9181951  
 R:Cast, M.J.; Sims, H.F.; Murdoch, G.L.; Gast, P.M.; Strauss, A.W.  
 Am. J. Obstet. Gynecol. 161, 1726-1731, 1989  
 A:Title: Isolation and sequencing of a complementary deoxyribonucleic acid clone encodin  
 A:Reference number: A60440; MUID:90102621  
 A:Accession: A60440  
 A:Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 6-68; 'G', 70-91, 'EAWAGAG', 102-158, 160-178, 'P', 181-311, 'A', 313-328 <GAS>  
 R:Murdoch, G.L.; Chin, C.C.; Offord, R.E.; Bradshaw, R.A.; Warren, J.C.  
 J. Biol. Chem. 258, 11460-11464, 1983  
 A:Title: Human placental estradiol 17-beta-dehydrogenase. Identification of a single his  
 A:Reference number: A21031; MUID:84008135  
 A:Accession: A21031  
 A:Molecule type: protein  
 A:Residues: 220-224 <MUR>  
 R:Murdoch, G.L.; Chin, C.C.; Warren, J.C.  
 Biochemistry 25, 641-646, 1986  
 A:Title: Human placental estradiol 17beta-dehydrogenase: sequence of a histidine-bearing  
 A:Reference number: A60628; MUID:86159758  
 A:Accession: A60628  
 A:Molecule type: protein  
 A:Residues: 208-215 <M02>  
 R:Labrie, F.; Luu-The, V.; Labrie, C.; Berube, D.; Couet, J.; Zhao, H.F.; Gagne, R.; Sim  
 J. Steroid Biochem. 34, 189-197, 1989  
 A:Title: Characterization of two mRNA species encoding human estradiol 17beta-dehydrogen  
 A:Reference number: A37190; MUID:90173172  
 A:Accession: A37190  
 A:Contents: annotation  
 R:Shen, Y.  
 submitted to the EMBL Data Library, August 1995  
 A:Reference number: G09150  
 A:Accession: G09150  
 A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA  
 A:Residues: 1-328 <SHE>  
 A:Cross-references: EMBL:U34879; NID:9375894; PIDN:AAD05019.1; PID:9375895  
 C:Genetics:  
 A:Gene: GDB:HS017B1; EDH17B2; EDHB  
 A:Cross-references: GDB:127970  
 A:Map position: 17q11-17q21  
 A:Introns: 33/1; 89/1; 149/1; 180/2; 239/3  
 C:Function:  
 A:Description: catalyzes the reversible oxidation of estradiol-17beta to estrone usin  
 A:Note: also catalyzes the reversible oxidation of 5-androstenediol to dehydroepiandr  
 C:Superfamily: estradiol 17beta-dehydrogenase; short-chain alcohol dehydrogenase homo  
 C:Keywords: oxidoreductase  
 E:2-328/Product: estradiol 17beta-dehydrogenase #status experimental <MAT>  
 F:4-189/Product: short-chain alcohol dehydrogenase homology <SADH>  
 F:211,214,222/Active site: His #status experimental

Query Match 61.4%; Score 35; DB 1; Length 328;  
 Best Local Similarity 50.0%; Pred. No. 20;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 PAPMKFFTT 11  
 Db 248 KPFTLXFTT 257

RESULT 13  
 G86263  
 hypothetical protein P3F9J.1 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001  
 C:Accession: G86263  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon  
 Chin, C.W.; Chung, M.K.; Corn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,  
 ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 815-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzia  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719  
 A:Accession: G86263  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-328 <STO>  
 A:Cross-references: GB:AE005172; NID:94850382; PIDN:AAD31052.1; GSPDB:GN00141  
 C:Genetics:  
 A:Map position: 1

Query Match 61.4%; Score 35; DB 2; Length 328;  
 Best Local Similarity 66.7%; Pred. No. 20;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PAPMKFFTT 10  
 Db 17 PNPTRKFTT 25

RESULT 14  
 T26360  
 hypothetical protein Y102A5C.21 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000  
 C:Accession: T26360  
 R:Gardner, A.  
 submitted to the EMBL Data Library, September 1998  
 A:Reference number: Z20204  
 A:Accession: T26360  
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
 A:Residues: 1-388 <WILL>  
 A:Cross-references: EMBL:AL031627; PIDN:CAA20962.1; CESP:Y102A5C.21  
 A:Experimental source: clone Y102A5C  
 C:Genetics:  
 A:Gene: CESP:Y102A5C.21  
 A:Introns: 154/1; 364/1  
 C:Superfamily: Caenorhabditis hypothetical protein C49G7.2

Query Match 61.4%; Score 35; DB 2; Length 388;  
 Best Local Similarity 60.0%; Pred. NO. 24;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 PAPMAKFFTT 11  
 | | : ||||  
 Db 220 PCPTREFFTT 229

## RESULT 15

H70216  
 PTS system, cellobiose-specific IIC component (celB) homolog - Lyme disease spirochete  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C>Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 21-Jul-2000  
 C:Accession: H70216  
 R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White  
 son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,  
 Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
 Nature 390, 580-586, 1997  
 A:Authors: Smith, H.O.; Venter, J.C.  
 A>Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.  
 A:Reference number: A70100; MUID:98065943  
 A:Accession: H70216  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-443 <RIE>  
 A:Cross-references: GB:AE00792; NID:93253098; PIDN:AAC66324.1; PID:92689900; TIGR:BBB04  
 A:Experimental source: strain B31  
 C:Genetics:  
 A:Genome: plasmid  
 C:Superfamily: phosphotransferase system enzyme II factor II, phosphoenolpyruvate-depend

Query Match 61.4%; Score 35; DB 2; Length 443;  
 Best Local Similarity 60.0%; Pred. NO. 28;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 PAPMAKFFTT 11  
 |||: ||:|  
 Db 396 PAPIAFFFFT 405

Search completed: June 11, 2002, 15:51:30  
 Job time: 221 sec

**This Page Blank (uspto)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: June 11, 2002, 15:46:47 ; Search time 13.03 seconds  
(without alignments)  
22.495 Million cell updates/sec

Title: US-09-727-198-5  
Perfect score: 57  
Sequence: 1 XPAPMKKFTTX 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database :

1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfilltest.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	61.4	327	1	US-08-375-962B-12
2	35	61.4	327	2	US-08-562-114B-12
3	35	61.4	327	4	US-08-729-594A-12
4	34	59.6	278	1	US-08-188-582-9
5	34	59.6	278	1	US-08-646-715-9
6	33	57.9	280	1	US-08-409-452-1
7	33	57.9	311	4	US-09-056-556-188
8	33	57.9	330	5	PCT-US93-08528-19
9	33	57.9	330	5	PCT-US93-08528-19
10	33	57.9	461	1	US-08-194-338-4
11	32	56.1	229	2	US-08-817-997A-2
12	32	56.1	371	4	US-09-500-569-16
13	32	56.1	1597	4	US-09-423-890-13
14	32	56.1	1597	4	US-08-628-829-14
15	31	54.4	376	4	US-09-500-569-10
16	31	54.4	481	6	5164481-1
17	30	52.6	275	1	US-08-252-995D-13
18	30	52.6	275	2	US-08-834-108-13
19	30	52.6	341	4	US-09-347-803-16
20	30	52.6	382	2	US-08-872-302-2
21	30	52.6	976	3	US-08-560-005-2
22	30	52.6	976	3	US-09-195-868-14
23	30	52.6	976	4	US-09-418-540-2
24	30	52.6	1187	4	US-08-664-962B-8
25	30	52.6	1187	4	US-09-311-743-8
26	30	52.6	1189	3	US-09-195-868-15
27	30	52.6	1229	3	US-09-195-868-28

28	29	50.9	56	1	US-08-358-160-87	Sequence 87, Appl
29	29	50.9	362	4	US-09-500-569-14	Sequence 14, Appl
30	29	50.9	769	3	US-09-320-878-12	Sequence 12, Appl
31	29	50.9	809	4	US-09-105-537-24	Sequence 24, Appl
32	29	50.9	824	2	US-08-785-310A-7	Sequence 7, Appl
33	29	50.9	824	2	US-08-816-693A-52	Sequence 52, Appl
34	29	50.9	824	3	US-08-885-291-52	Sequence 52, Appl
35	29	50.9	824	4	US-09-496-672-52	Sequence 52, Appl
36	29	50.9	1093	3	US-08-545-860D-55	Sequence 55, Appl
37	29	50.9	1093	5	PCT-US94-04496-55	Sequence 55, Appl
38	29	50.9	1185	4	US-09-041-886-23	Sequence 23, Appl
39	29	50.9	3782	4	US-09-105-537-4	Sequence 4, Appl
40	28	49.1	123	4	US-09-500-569-28	Sequence 28, Appl
41	28	49.1	129	4	US-08-942-686-3	Sequence 3, Appl
42	28	49.1	232	1	US-07-797-556-4	Sequence 4, Appl
43	28	49.1	232	1	US-08-225-989-4	Sequence 4, Appl
44	28	49.1	232	1	US-08-570-923-4	Sequence 4, Appl
45	28	49.1	232	1	US-08-580-014-4	Sequence 4, Appl

## ALIGNMENTS

RESULT 1  
US-08-375-962B-12  
Sequence 12, Application US/08375962B  
Patent No. 5731195  
GENERAL INFORMATION:  
APPLICANT: SIMON, ANDRAS; HELLMAN, ULF; WERNSTEDT,  
APPLICANT: CHRISTER, ERIKSSON, ULF.  
TITLE OF INVENTION: Isolated Nucleic Acid Molecule  
WHICH CODES FOR A 32 kDa Protein Having 11-Cis Retinol  
TITLE OF INVENTION: Dehydrogenase Activity, and Which Associates With p63,  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Felife & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect (ASCII standard)  
CURRENT APPLICATION DATA: US/08375,962B  
APPLICATION NUMBER: US/08375,962B  
FILING DATE: 20-January-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/258,418  
FILING DATE: 6-October-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Pasqualini, Patricia A.  
REGISTRATION NUMBER: 34,894  
REFERENCE/DOCKET NUMBER: LUD 5372.1 CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 327 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Human estradiol 17-b dehydrogenase (hEDH)  
US-08-375-962B-12  
Query Match 61.4%, Score 35; DB 1; Length 327;

Best Local Similarity 50.0%; Pred. No. 20;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 PAPMKFFTT 11  
| | : : : | | |  
Db 247 PKPTLRFTT 256

RESULT 2  
US-08-562-114B-12

; Sequence 12, Application US/08562114B  
; Patent No. 5972646

; GENERAL INFORMATION:

; APPLICANT: ERIKSSON ET AL.

; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE WHICH CODES FOR A

; TITLE OF INVENTION: 32 KDA PROTEIN HAVING 11-CIS RETINOL HYDROGENASE

; TITLE OF INVENTION: ACTIVITY, AND WHICH ASSOCIATES WITH P63, A

; TITLE OF INVENTION: PORTION OF A RETINOL BINDING PROTEIN RECEPTOR

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felt & Lynch

; STREET: 805 Third Avenue

; CITY: New York City

; STATE: New York

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage

; COMPUTER: IBM

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: Wordperfect 5.1 and ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/562,114B

; FILING DATE: 22-No. 5972646ember-1995

; CLASSIFICATION: 435

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 08/375,962

; FILING DATE: 20-January-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Kohl, Vineet

; REGISTRATION NUMBER: 37,003

; REFERENCE/DOCKET NUMBER: LUD 5372.2 CIP

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 688-9200

; TELEFAX: (212) 838-3884

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 327 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; FEATURE:

; NAME/KEY: Human estradiol 17-b dehydrogenase (hbdh)

; US-08-562-114B-12

Query Match 61.4%; Score 35; DB 2; Length 327;

Best Local Similarity 50.0%; Pred. No. 20;

Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 PAPMKFFTT 11  
| | : : : | | |  
Db 247 PKPTLRFTT 256

RESULT 3

US-08-729-594A-12

; Sequence 12, Application US/08729594A

; Patent No. 6280997

; GENERAL INFORMATION:

; APPLICANT: Eriksson, Ulf; Simon, Andreas; Romert, Anna

; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE WHICH

; TITLE OF INVENTION: CODES FOR A 32 KDA PROTEIN HAVING 11-CIS RETINOL DEHYDROGENASE

; TITLE OF INVENTION: ACTIVITY, AND WHICH ASSOCIATES WITH P63, A PORTION OF A

; TITLE OF INVENTION: RETINOL BINDING PROTEIN RECEPTOR

; NUMBER OF SEQUENCES: 41

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felt & Lynch

; STREET: 805 Third Avenue

; CITY: New York City

; STATE: New York

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage

; COMPUTER: IBM

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/729,594A

; FILING DATE: 11-October-1996

; CLASSIFICATION: 435

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 08/562,114

; FILING DATE: 22-No. 6280997ember-1995

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 08/258,418

; FILING DATE: 10-June-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Hanson, No. 6280997man D.

; REGISTRATION NUMBER: 30,946

; REFERENCE/DOCKET NUMBER: LUD 5372.3

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 688-9200

; TELEFAX: (212) 838-3884

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 327 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; FEATURE:

; NAME/KEY: Human estradiol 17-b dehydrogenase (hbdh)

; US-08-729-594A-12

Query Match 61.4%; Score 35; DB 4; Length 327;

Best Local Similarity 50.0%; Pred. No. 20;

Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 PAPMKFFTT 11  
| | : : : | | |  
Db 247 PKPTLRFTT 256

RESULT 4

US-08-188-582-9

; Sequence 9, Application US/08188582

; Patent No. 5534410

; GENERAL INFORMATION:

; APPLICANT: Tjian, Robert

; APPLICANT: Comal, Lucio

; APPLICANT: Dynlacht, Brian D.

; APPLICANT: Hoey, Timothy

; APPLICANT: Ruppert, Siegfried

; APPLICANT: Tanese, Naoko

; APPLICANT: Weinzierl, Robert O.J.

; TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,

; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE

; NUMBER OF SEQUENCES: 36

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT

; STREET: 4 Embarcadero Center, Suite 3400

; CITY: San Francisco

STATE: California  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/188,582  
FILING DATE: 28-JAN-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman, Richard A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 278 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-188-582-9

Query Match 59.6%; Score 34; DB 1; Length 278;  
Best Local Similarity 60.0%; Pred. No. 27;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PAPMKFFTT 11  
11:111111  
Db 183 PKVVKFTTT 192

RESULT 5  
US-08-646-715-9  
Sequence 9, Application US/08646715  
Patent No. 5637686  
GENERAL INFORMATION:  
APPLICANT: Tjian, Robert  
APPLICANT: Comai, Lucio  
APPLICANT: Dynlacht, Brian D.  
APPLICANT: Hoey, Timothy  
APPLICANT: Ruppert, Siegfried  
APPLICANT: Tanese, Naoko  
APPLICANT: Wang, Edith  
APPLICANT: Weinzierl, Robert O.J.  
TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,  
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,715  
FILING DATE: 09-MAY-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/188,582  
FILING DATE: 28-JAN-1994

ATTORNEY/AGENT INFORMATION:  
NAME: Osman, Richard A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 278 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-646-715-9

Query Match 59.6%; Score 34; DB 1; Length 278;  
Best Local Similarity 60.0%; Pred. No. 27;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PAPMKFFTT 11  
11:111111  
Db 183 PKVVKFTTT 192

RESULT 6  
US-08-409-452-1  
Sequence 1, Application US/08409452  
Patent No. 5747293  
GENERAL INFORMATION:  
APPLICANT: DOUGAN, Gordon  
APPLICANT: FRANKEL, Gad  
TITLE OF INVENTION: Antibodies to Intimin-like  
TITLE OF INVENTION: Proteins of E. coli  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: HALE and DORR  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: DISTRICT OF COLUMBIA  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/409,452  
FILING DATE: 23-MAR-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: BAKER, Hollie L.  
REGISTRATION NUMBER: 31,321  
REFERENCE/DOCKET NUMBER: 102286,319  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 942-8400  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 280 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-409-452-1

Query Match 57.9%; Score 33; DB 1; Length 280;  
Best Local Similarity 66.7%; Pred. No. 42;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 APMKFFTT 11

Db 86 APEVEFTT 94

RESULT 7  
US-09-056-556-188  
; Sequence 188, Application US/09056556  
; Patent No. 6350456  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, David C.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND  
; NUMBER OF SEQUENCES: 241  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED AND BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/056,556  
; FILING DATE: 07-APR-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.457  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 683-6031  
; INFORMATION FOR SEQ ID NO: 188:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 311 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; US-09-056-556-188

Query Match 57.9%; Score 33; DB 4; Length 311;  
Best Local Similarity 50.0%; Pred. No. 47;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 PAPMAKFTT 11  
Db 8 PVPMEFLTS 17

RESULT 8  
US-08-118-270-19  
; Sequence 19, Application US/08118270  
; Patent No. 5508384  
; GENERAL INFORMATION:  
; APPLICANT: Murphy, Randall B.  
; APPLICANT: Schuster, David I.  
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN  
; NUMBER OF SEQUENCES: 348  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMARK  
; STREET: 419 Seventh Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; SEQUENCE CHARACTERISTICS:

MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/118,270  
; FILING DATE: 09-SEP-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/943,236  
; FILING DATE: 10-SEP-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Townsend, Kevin G.  
; REGISTRATION NUMBER: 34,033  
; REFERENCE/DOCKET NUMBER: MURPHY-2A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-628-5197  
; TELEFAX: 202-737-3528  
; TELEEX: 248633  
; INFORMATION FOR SEQ ID NO: 19:  
; MOLECULE TYPE: peptide  
; US-08-118-270-19

Query Match 57.9%; Score 33; DB 1; Length 330;  
Best Local Similarity 62.5%; Pred. No. 50;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PAPMAKEF 9  
Db 283 PGPLFKFF 290

RESULT 9  
PCT-US93-08528-19  
; Sequence 19, Application PC/TUS9308528  
; GENERAL INFORMATION:  
; APPLICANT: New York University  
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN  
; NUMBER OF SEQUENCES: 348  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMARK  
; STREET: 419 Seventh Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/08528  
; FILING DATE: 09-SEP-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/943,236  
; FILING DATE: 10-SEP-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Townsend, Kevin G.  
; REGISTRATION NUMBER: 34,033  
; REFERENCE/DOCKET NUMBER: MURPHY-2 PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-628-5197  
; TELEFAX: 202-737-3528  
; TELEEX: 248633  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:

LENGTH: 330 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US93-08528-19

Query Match 57.9%; Score 33; DB 5; Length 330;  
Best Local Similarity 62.5%; Pred. No. 50;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 PAPMMKFF 9  
1 1 1 1 1  
DB 283 PGPFLKFF 290

RESULT 10  
US-08-194-338-4  
Sequence 4, Application US/08194338  
Patent No. 5474898  
GENERAL INFORMATION:  
APPLICANT: Venter, John C.  
APPLICANT: Fraser, Claire M.  
APPLICANT: McCormick, William R.  
TITLE OF INVENTION: OCTOPAMINE RECEPTOR  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobb, Martens, Olson and Bear  
STREET: 620 Newport Center Drive, Sixteenth Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/194,338  
FILING DATE: 08-FEB-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/676,174  
FILING DATE: 28-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelson, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: NIH101.001DV1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 461 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
US-08-194-338-4

Query Match 57.9%; Score 33; DB 1; Length 461;  
Best Local Similarity 62.5%; Pred. No. 71;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
OY 2 PAPMMKFF 9  
1 1 1 1 1  
DB 414 PGPFLKFF 421

RESULT 11  
US-08-817-997A-2  
Sequence 2, Application US/08817997A  
Patent No. 5827719  
GENERAL INFORMATION:  
APPLICANT: Sandel, Thomas  
APPLICANT: Kauppinen, Sakari  
APPLICANT: Kofod, Lene V.  
TITLE OF INVENTION: An Enzyme with Lipolytic  
TITLE OF INVENTION: Activity  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 58277190 No. 5827719disk Of No. 5827719th America, Inc.  
STREET: 405 Lexington Avenue - 64th Fl.  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10174  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/817,997A  
FILING DATE: 24-APR-1997  
CLASSIFICATION: 7435  
ATTORNEY/AGENT INFORMATION:  
NAME: Lambiris, Elias  
REGISTRATION NUMBER: 33,728  
REFERENCE/DOCKET NUMBER: 4316.204  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-878-9652  
TELEFAX: 212-878-9653  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 229 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: Internal  
US-08-817-997A-2

Query Match 56.1%; Score 32; DB 2; Length 229;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 MKFTT 11  
1 1 1 1 1 1  
DB 1 MKFTT 6

RESULT 12  
US-09-500-569-16  
Sequence 16, Application US/09500569  
Patent No. 6329204  
GENERAL INFORMATION:  
APPLICANT: Cahoon, Rebecca E.  
APPLICANT: Rafalski, Antoni  
APPLICANT: Shen, Jennie  
TITLE OF INVENTION: Plant Caffeic acid 3-O-Methyltransferase Homologs  
FILE REFERENCE: BB1327 US NA  
CURRENT APPLICATION NUMBER: US/09/500,569  
CURRENT FILING DATE: 2000-02-09  
EARLIER APPLICATION NUMBER: 60/119,587  
EARLIER FILING DATE: 1999-February-10  
NUMBER OF SEQ ID NOS: 28  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 16

; LENGTH: 371  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-09-500-569-16

Query Match 56.1%; Score 32; DB 4; Length 371;  
Best Local Similarity 55.6%; Pred. No. 88;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 PAPMKFFT 10  
|||: |||  
Db 114 PAPVCRWFT 122

RESULT 13  
US-09-423-890-13  
; Sequence 13, Application US/09423890  
; Patent No. 6312934  
; GENERAL INFORMATION:

; APPLICANT: CADUS PHARMACEUTICAL CORPORATION  
; TITLE OF INVENTION: HUMAN MEK PROTEIN AND NUCLEIC ACID MOLECULES  
; FILE REFERENCE: CPI-085CPC  
; CURRENT APPLICATION NUMBER: US/09/423,890

; PRIOR FILING DATE: 2000-03-06  
; PRIOR APPLICATION NUMBER: USSN 60/078,153  
; PRIOR FILING DATE: 1998-03-16  
; PRIOR APPLICATION NUMBER: USSN 60/099,165  
; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 1597  
; TYPE: PRT  
; ORGANISM: Mus musculus

US-09-423-890-13

Query Match 56.1%; Score 32; DB 4; Length 1597;  
Best Local Similarity 60.0%; Pred. No. 4.1e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 PAPMKFFT 11  
|||: |||  
Db 212 PADRLKFFT 221

RESULT 14  
US-08-628-829-14

; Sequence 14, Application US/08628829A  
; Patent No. 6333170  
; GENERAL INFORMATION:

; APPLICANT: Johnson, Gary L.  
; TITLE OF INVENTION: Method And Product For Regulating Cell Responsiveness To External  
; FILE REFERENCE: CPI-004DVC3  
; CURRENT APPLICATION NUMBER: US/08/628,829A  
; PRIOR FILING DATE: 1996-04-05  
; PRIOR APPLICATION NUMBER: 08/440,421  
; PRIOR FILING DATE: 1995-05-15  
; PRIOR APPLICATION NUMBER: 08/323,460  
; PRIOR FILING DATE: 1994-10-14  
; PRIOR APPLICATION NUMBER: 08/049,254  
; PRIOR FILING DATE: 1993-05-15  
; PRIOR APPLICATION NUMBER: 08/410,602  
; PRIOR FILING DATE: 1995-04-24  
; PRIOR APPLICATION NUMBER: 08/472,934  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 14  
; LENGTH: 1597  
; TYPE: PRT

; ORGANISM: Mus musculus  
US-08-628-829-14

Query Match 56.1%; Score 32; DB 4; Length 1597;  
Best Local Similarity 60.0%; Pred. No. 4.1e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 PAPMKFFT 11  
|||: |||  
Db 212 PADRLKFFT 221

RESULT 15  
US-09-500-569-10  
; Sequence 10, Application US/09500569  
; Patent No. 6329204  
; GENERAL INFORMATION:

; APPLICANT: Cahoon, Rebecca E.  
; APPLICANT: Rafalski, Antoni

; TITLE OF INVENTION: Plant Caffeic acid 3-O-Methyltransferase Homologs  
; FILE REFERENCE: BR1327 US NA  
; CURRENT APPLICATION NUMBER: US/09/500,569  
; PRIOR FILING DATE: 2000-02-09  
; PRIOR APPLICATION NUMBER: 60/119,587  
; PRIOR FILING DATE: 1999-February-10  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 10  
; LENGTH: 378  
; TYPE: PRT  
; ORGANISM: Oryza sativa

US-09-500-569-10

Query Match 54.4%; Score 31; DB 4; Length 378;  
Best Local Similarity 55.6%; Pred. No. 1.4e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 PAPMKFFT 10  
|||: |||  
Db 118 PAPVCRWFT 126

Search completed: June 11, 2002, 15:51:20  
Job time: 273 sec



**This Page Blank (uspto)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd

OM protein - protein search, using sw model

Run on: June 11, 2002, 15:36:21 ; Search time 30.08 Seconds  
(without alignments)  
44.311 Million cell updates/sec

```
Title: US-09-727-198-5
Perfect score: 57
Sequence: 1 XPAPMMKFFTTX 12
```

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

```
Searched:      747574 segs, 111073796 residues
Total number of hits satisfying chosen parameters: 747574
```

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

Database : A.geneseq\_037802.\*

1:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT.*
2:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT.*
3:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT.*
4:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT.*
5:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT.*
6:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT.*
7:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT.*
8:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT.*
9:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT.*
10:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT.*
11:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT.*
12:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT.*
13:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT.*
14:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT.*
15:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT.*
16:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT.*
17:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT.*
18:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT.*
19:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT.*
20:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT.*
21:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT.*
22:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	Score	Query Match	Length	DB	ID	Description
No.						
1	45	78.9	1332	22	ABG09133	Novel human diatom
2	36	63.2	19	22	ABB30037	Peptide #2688 encoc
3	36	63.2	19	22	ABB35209	Peptide #2715 encoc
4	36	63.2	19	22	ABB20650	Protein #2649 encoc
5	36	63.2	19	22	AAH56041	Human brain expres
6	36	63.2	19	22	AAH68412	Human bone marrow
7	36	63.2	19	22	AAH16227	Peptide #2661 encoc
8	36	63.2	19	22	AAH28717	Peptide #2754 encoc
9	36	63.2	19	22	AAH03957	Peptide #2639 encoc
10	36	63.2	140-	21	AAH12771	Arabidopsis thailia
11	35	61.4	2016	22	ABH63911	Drosophila melanog

45	56.1	189	21	AAG33435	
44	56.1	123	21	ABG03943	
43	102	21		AAH40892	
42	56.1	100	22	AAU44811	
41	57.9	1338	22	ABB66084	
40	57.9	1338	22	ABB71863	
39	57.9	640	22	ABB62091	
38	57.9	491	22	ABB70144	
37	57.9	475	22	ABB67155	
36	57.9	461	22	AAH54834	
35	57.9	458	15	AAH54834	
34	57.9	457	22	AAH52126	
33	57.9	362	9	AAH81100	
32	57.9	330	17	AAH02670	
31	57.9	330	15	AAH48698	
30	57.9	311	20	AAH39168	
29	57.9	311	19	AAH64371	
28	57.9	311	19	AAH81736	
27	57.9	280	21	AAU06699	
26	57.9	280	22	AAU06699	
25	57.9	280	22	AAU06699	
24	57.9	280	22	AAU06699	
23	57.9	280	22	AAU06699	
22	57.9	280	22	AAU06699	
21	57.9	280	22	AAU06699	
20	57.9	280	22	AAU06699	
19	57.9	280	22	AAU06699	
18	57.9	280	22	AAU06699	
17	57.9	280	22	AAU06699	
16	57.9	280	22	AAU06699	
15	57.9	280	22	AAU06699	
14	57.9	280	22	AAU06699	
13	57.9	280	22	AAU06699	
12	57.9	280	22	AAU06699	
11	57.9	280	22	AAU06699	
10	57.9	280	22	AAU06699	
9	57.9	280	22	AAU06699	
8	57.9	280	22	AAU06699	
7	57.9	280	22	AAU06699	
6	57.9	280	22	AAU06699	
5	57.9	280	22	AAU06699	
4	57.9	280	22	AAU06699	
3	57.9	280	22	AAU06699	
2	57.9	280	22	AAU06699	
1	57.9	280	22	AAU06699	

## ALIGNMENTS

	RESULT	1
ABG09133	ID	ABG09133 standard; Protein; 1332 AA.
XX	AC	ABG09133;
XX	DT	13-FEB-2002 (first entry)
XX	DE	Novel human diagnostic protein #9124.
KW	KW	Human: chromosome mapping; gene therapy; forensic;
KW	KW	food supplement; medical imaging; diagnostic; genetic disorder.
OS	OS	Homo sapiens.
PN	PN	WO200175067-A2.
PD	PD	11-OCT-2001.
PF	PF	30-MAR-2001; 2001MO-US08631.
PR	PR	31-MAR-2000; 2000US-0540217.
PA	PA	23-AUG-2000; 2000US-0649167.
XX	XX	(HYSE-) HYSEQ INC.
PI	PI	Dimanac RT, Liu C, Tang YT;
DR	DR	WPI: 2001-639362/73.
DR	DR	N-PDB: AAS73320.
FT	FT	New isolated polynucleotide and encoded polypeptides, useful in
FT	FT	diagnostics, forensics, gene mapping, identification of mutations
FT	FT	responsible for genetic disorders or other traits and to assess

PT biodiversity -  
XX  
PS Claim 20; SEQ ID No 39492; 103bp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
SQ Sequence 1332 AA;

Query Match 78.9%; Score 45; DB 22; Length 1332;  
Best Local Similarity 80.0%; Pred. No. 4;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 PAPMKFFTT 11  
| | | : | | | |  
Db 76 papmkfftt 85

## RESULT 2

ABR30037 standard; Peptide; 19 AA.

AC ABR30037;

DT 01-FEB-2002 (first entry)

DE Peptide #2688 encoded by breast cell single exon nucleic acid probe.

XX Human; microarray; single exon probe; gene expression; breast;

KM disease; cancer.

XX Homo sapiens.

PN WO200157271-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00062.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-496933/54.

XX

PT New spatially-addressable set of single exon nucleic acid probes,  
PT useful for measuring gene expression in sample derived from human  
PT breast, comprises number of single exon nucleic acid probes -  
XX  
PS Claim 27; SEQ ID NO 13005; 327bp + sequence listing; English.  
XX

CC The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human breast and BR 474 cells. The method involves contacting  
CC the probes with a collection of detectably labelled nucleic acids  
CC derived from mRNA of human breast, and then measuring the label  
CC bound to each probe of the microarray. The probes are useful for  
CC verifying the expression of regions of genomic DNA predicted to  
CC encode proteins. They are useful for gene discovery, and for  
CC determining predisposition and/or prognosing breast disease. Gene  
CC expression analysis is useful for assessing the toxicity of chemical  
CC agents on cells. The microarray of this invention presents a far greater  
CC diversity of probes for measuring gene expression, with far less bias  
CC than expressed sequence tag microarrays. The method is suitable for  
CC rapid production of functional information from genomic sequence. The  
CC present sequence is a peptide encoded by a single exon nucleic acid  
CC probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 19 AA;

Query Match 63.2%; Score 36; DB 22; Length 19;  
Best Local Similarity 66.7%; Pred. No. 3.3;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 PAPMKFFTT 10  
| | | | |  
Db 2 prpqmsfftt 10

## RESULT 3

ABR35209 standard; Peptide; 19 AA.

AC ABR35209;

DT 04-FEB-2002 (first entry)

DE Peptide #2715 encoded by human foetal liver single exon probe.

XX Human; foetal liver; gene expression; single exon nucleic acid probe.

KM Homo sapiens.

PN WO200157277-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00069.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483447/52.

XX

Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human fetal liver -  
XX  
PS Claim 27; SEQ ID NO 27844; 639bp + sequence listing; English.  
XX  
CC The invention relates to a single exon nucleic acid probe for  
CC measuring human gene expression in a sample derived from human foetal  
CC liver. The single exon nucleic acid probes may be used for predicting,  
CC measuring and displaying gene expression in samples derived from human  
CC fetal liver. The present sequence is a peptide encoded by a single exon  
CC nucleic acid probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 19 AA:  
  
Query Match 63.2%; Score 36; DB 22; Length 19;  
Best Local Similarity 66.7%; Pred. No. 3.3;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
OY 2 PAPMKKFFT 10  
Db 2 ptpqmsfft 10  
  
RESULT 4  
ID ABB20650 standard; Protein: 19 AA.  
XX  
AC ABB20650;  
XX  
DT 23-JAN-2002 (first entry)  
XX  
DE Protein #2649 encoded by probe for measuring heart cell gene expression.  
XX  
KW Human; gene expression: heart; microarray; vascular system;  
KW cardiovascular disease; hypertension; cardiac arrhythmia;  
KW congenital heart disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200157274-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00666.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488899/53.  
XX  
PT Single exon nucleic acid probes for analyzing gene expression in human  
PT hearts -  
XX  
PS Claim 15; SEQ ID NO 22420; 530pp; English.  
XX  
CC The present invention relates to single exon nucleic acid probes for  
CC measuring human gene expression in a sample derived from human heart (see  
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
CC probe. The probes may be used for predicting, measuring and displaying  
CC gene expression in samples derived from the human heart via microarrays.  
CC By measuring gene expression, the probes are useful for predicting.

CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
CC human heart and vascular system e.g. cardiovascular disease,  
CC hypertension, cardiac arrhythmias and congenital heart disease.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 19 AA:  
  
Query Match 63.2%; Score 36; DB 22; Length 19;  
Best Local Similarity 66.7%; Pred. No. 3.3;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
OY 2 PAPMKKFFT 10  
Db 2 ptpqmsfft 10  
  
RESULT 5  
ID AAM56041 standard; Protein: 19 AA.  
XX  
AC AAM56041;  
XX  
DT 05-NOV-2001 (first entry)  
XX  
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 28146.  
XX  
KW Human; brain expressed exon; gene expression analysis; probe;  
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
KW epilepsy; cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO200157275-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00667.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-483446/52.  
XX  
PT Single exon nucleic acid probes for analyzing gene expression in human  
PT brains -  
XX  
PS Example 4; SEQ ID NO: 28146; 650pp + Sequence Listing; English.  
XX  
CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC brain. They can be used to measure gene expression in brain cell samples,  
CC which may enable the diagnosis and improved treatment of nervous system  
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
CC epilepsy and cancers. The present sequence is a protein encoded by one of  
CC the probes of the invention.  
XX  
SQ Sequence 19 AA:  
  
Query Match 63.2%; Score 36; DB 22; Length 19;  
Best Local Similarity 66.7%; Pred. No. 3.3;

```

Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 2 PAPMKFFT 10
   | | | | |
   2 ptpqmsfft 10
Db
RESULT 6
AAM68412
ID AAM68412 standard; Protein; 19 AA.
XX
AC AAM68412;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 28718.
XX
KM Human; bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 28718; 658bp + Sequence Listing; English.
XX
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
SQ Sequence 19 AA;

```

```

Query Match 63.2%; Score 36; DB 22; Length 19;
Best Local Similarity 66.7%; Pred. No. 3.3;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

OY 2 PAPMKFFT 10
   | | | | |
   2 ptpqmsfft 10
Db

```

```

RESULT 7
AAM16227
ID AAM16227 standard; Protein; 19 AA.
XX
AC AAM16227;
XX
OS Homo sapiens.
XX
DT 12-OCT-2001 (first entry)

```

```

XX
DE Peptide #2661 encoded by probe for measuring cervical gene expression.
XX
KM Probe: human; microarray; gene expression; cervical epithelial cell;
XX cervical cancer.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
PS Claim 27; SEQ ID No 21053; 487bp; English.
XX
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP: see AAI10068-AA128459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 19 AA;

```

```

Query Match 63.2%; Score 36; DB 22; Length 19;
Best Local Similarity 66.7%; Pred. No. 3.3;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

OY 2 PAPMKFFT 10
   | | | | |
   2 ptpqmsfft 10
Db

```

```

RESULT 8
AAM28717
ID AAM28717 standard; Protein; 19 AA.
XX
AC AAM28717;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #2754 encoded by probe for measuring placental gene expression.
XX
KM Probe: microarray; human; placenta; antenatal diagnosis;
XX genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.

```

XX 09-AUG-2001.  
PD Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI: 2001-488897/53.  
PF 30-JAN-2001; 2001WO-US00663.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PT 26-MAY-2000; 2000US-0207456.  
XX 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
XX 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
XX 04-OCT-2000; 2000GB-0024263.  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI: 2001-488897/53.  
DR  
XX Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human placenta -  
XX  
PS Claim 27; SEQ ID No 28986; 654bp; English.  
XX  
CC The present invention relates to single exon nucleic acid probes (SENPs;  
CC see AA13115-AA157346). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for producing a microarray for  
CC predicting, measuring and displaying gene expression in samples derived  
CC from human placenta. The probes are useful for antenatal diagnosis of  
CC human genetic disorders.  
XX  
SQ Sequence 19 AA:  
  
Query Match 63.2%; Score 36; DB 22; Length 19;  
Best Local Similarity 66.7%; Pred. No. 3.3;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 2 PAPMMKFFP 10  
DB 2 ptpqmsfft 10  
| | | | |  
| | | | |  
  
RESULT 9  
AA03957  
ID AA03957 standard; Protein; 19 AA.  
XX  
AC AA03957;  
XX  
DT 09-OCT-2001 (first entry)  
XX  
DE Peptide #2639 encoded by probe for measuring breast gene expression.  
XX  
KW Probe; human; breast disease; breast cancer; development disorder;  
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.  
XX  
OS Homo sapiens.  
XX  
PN WO200157270-A2.  
PD  
XX 09-AUG-2001.  
PF 29-JAN-2001; 2001WO-US00661.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PT 26-MAY-2000; 2000US-0207456.  
XX 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
XX 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
XX 04-OCT-2000; 2000GB-0024263.  
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI: 2001-476286/51.  
DR  
XX Novel single exon nucleic acid probe used to measuring gene expression  
PT in a human breast -  
XX  
PS Claim 27; SEQ ID No 12697; 322bp; English.  
XX  
CC The present invention relates to novel single exon nucleic acid probes  
CC (see AA100010-AA110067). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for measuring human gene expression in  
CC a human breast sample, where the probe hybridises at high stringency to a  
CC nucleic acid expressed in the human breast. The probes are useful for  
CC predicting, diagnosing, grading, staging, monitoring and prognosis  
CC diseases of the human breast, particularly those diseases with polygenic  
CC aetiology. The diseases include: breast cancer, disorders of development,  
CC inflammatory diseases of the breast, fibrocystic changes, proliferative  
CC breast disease and non-carcinoma tumours.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 19 AA:  
  
Query Match 63.2%; Score 36; DB 22; Length 19;  
Best Local Similarity 66.7%; Pred. No. 3.3;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 2 PAPMMKFFP 10  
DB 2 ptpqmsfft 10  
| | | | |  
| | | | |  
  
RESULT 10  
AAG12771  
ID AAG12771 standard; Protein; 140 AA.  
XX  
AC AAG12771;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SPQ ID NO: 12012.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
PD  
XX 06-SEP-2000.  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PT 05-MAR-1999; 99US-0123180.  
XX 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
XX 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
XX 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
XX 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
XX 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
XX 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
XX 28-APR-1999; 99US-0131449.

PR	30-APR-1999;	99US-0132048.
PR	30-APR-1999;	99US-0132407.
PR	04-MAY-1999;	99US-0132484.
PR	05-MAY-1999;	99US-0132485.
PR	06-MAY-1999;	99US-0132486.
PR	07-MAY-1999;	99US-0132487.
PR	11-MAY-1999;	99US-0132486.
PR	14-MAY-1999;	99US-0134218.
PR	14-MAY-1999;	99US-0134219.
PR	14-MAY-1999;	99US-0134221.
PR	18-MAY-1999;	99US-0134370.
PR	18-MAY-1999;	99US-0134768.
PR	20-MAY-1999;	99US-0134941.
PR	21-MAY-1999;	99US-0135124.
PR	24-MAY-1999;	99US-0135353.
PR	25-MAY-1999;	99US-0135629.
PR	27-MAY-1999;	99US-0136021.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137502.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	14-JUN-1999;	99US-0138847.
PR	16-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	16-JUN-1999;	99US-0139453.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143544.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145088.
PR	21-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	28-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	05-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147199.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148563.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18	

PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

## Query Match

Best Local Similarity 66.7%; Score 36; DB 21; Length 140;  
 Pred. No. 23;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 APMKFFTT 11  
 :|:|||||  
 Db 27 spykfftt 35

## RESULT 11

ABR63911  
 ID ABR63911 standard; Protein; 2016 AA.

XX ABR63911;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 18525.

XX Drosophila: developmental biology; cell signalling; insecticide;

KM pharmaceutical.

XX Drosophila melanogaster.

OS WO200171042-A2.

XX 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

PI WPI; 2001-656860/75.

DR N-PSDB; ABL08014.

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -

XX Disclosure; SEQ ID NO 18525; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABR63911), expressed DNA  
 CC sequences (ABR63911) and the encoded proteins

CC (ABR63911-ABR63911).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 2016 AA;

## Query Match

Best Local Similarity 61.4%; Score 35; DB 22; Length 2016;  
 Pred. No. 4.7e+02;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 PAPMKFFTT 11  
 :|:|||||  
 Db 1118 pspqsfst 1127

## RESULT 12

AA002149  
 ID AA002149 standard; Protein; 145 AA.

XX AA002149;

XX 06-NOV-2001 (first entry)

DE Human polypeptide SEQ ID NO 16041.

XX Human: cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukemia;

KW nervous system disorders; arthritis; inflammation.

XX Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

PF 26-FEB-2001; 2001WO-US04927.

XX 28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RF;

PI WPI; 2001-514838/56.

DR N-PSDB; AAI82080.

PT Isolated nucleic acids and polypeptides, useful for preventing  
 PT diagnosing and treating e.g. leukemia, inflammation and immune  
 PT disorders -

XX Claim 20; SEQ ID NO 16041; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and  
 CC the encoded proteins (AA000010-AA013910) that exhibit activity relating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activity/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukemia, nervous system disorders, arthritis and  
 CC inflammation.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 145 AA;

Query Match 59.6%; Score 34; DB 22; Length 145;  
Best Local Similarity 66.7%; Pred. No. 56;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 2 PAPMMKFFT 10  
I I I I I I  
Db 25 PPPMkffs 33

RESULT 13

AAG38411  
ID AAG38411 standard; Protein: 155 AA.

XX AAG38411;

DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 47383.

XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

PF 25-FEB-1999; 990S-0121825.

PR 05-MAR-1999; 990S-0123180.

PR 09-MAR-1999; 990S-0123548.

PR 23-MAR-1999; 990S-0125788.

PR 25-MAR-1999; 990S-0126264.

PR 29-MAR-1999; 990S-0126785.

PR 01-APR-1999; 990S-0127462.

PR 06-APR-1999; 990S-0128234.

PR 08-APR-1999; 990S-0128714.

PR 16-APR-1999; 990S-0129845.

PR 19-APR-1999; 990S-0130077.

PR 21-APR-1999; 990S-0130449.

PR 23-APR-1999; 990S-0130510.

PR 28-APR-1999; 990S-0130891.

PR 30-APR-1999; 990S-0132048.

PR 04-MAY-1999; 990S-0132484.

PR 05-MAY-1999; 990S-0132485.

PR 06-MAY-1999; 990S-0132486.

PR 07-MAY-1999; 990S-0132487.

PR 11-MAY-1999; 990S-0132863.

PR 14-MAY-1999; 990S-0134218.

PR 14-MAY-1999; 990S-0134219.

PR 14-MAY-1999; 990S-0134221.

PR 14-MAY-1999; 990S-0134370.

PR 10-JUN-1999; 990S-0138847.  
PR 14-JUN-1999; 990S-0139119.  
PR 16-JUN-1999; 990S-0139452.  
PR 16-JUN-1999; 990S-0139453.  
PR 17-JUN-1999; 990S-0139492.  
PR 18-JUN-1999; 990S-0139454.  
PR 18-JUN-1999; 990S-0139455.  
PR 18-JUN-1999; 990S-0139456.  
PR 18-JUN-1999; 990S-0139457.  
PR 18-JUN-1999; 990S-0139458.  
PR 18-JUN-1999; 990S-0139459.  
PR 18-JUN-1999; 990S-0139460.  
PR 18-JUN-1999; 990S-0139461.  
PR 18-JUN-1999; 990S-0139462.  
PR 18-JUN-1999; 990S-0139463.  
PR 18-JUN-1999; 990S-0139750.  
PR 18-JUN-1999; 990S-0139763.  
PR 21-JUN-1999; 990S-0139817.  
PR 22-JUN-1999; 990S-0139899.  
PR 23-JUN-1999; 990S-0140353.  
PR 23-JUN-1999; 990S-0140354.  
PR 24-JUN-1999; 990S-0140695.  
PR 28-JUN-1999; 990S-0140823.  
PR 28-JUN-1999; 990S-0140991.  
PR 30-JUN-1999; 990S-0141287.  
PR 01-JUL-1999; 990S-0141842.  
PR 01-JUL-1999; 990S-0142154.  
PR 02-JUL-1999; 990S-0142055.  
PR 06-JUL-1999; 990S-0142390.  
PR 08-JUL-1999; 990S-0142803.  
PR 09-JUL-1999; 990S-0142920.  
PR 12-JUL-1999; 990S-0142977.  
PR 13-JUL-1999; 990S-0143542.  
PR 14-JUL-1999; 990S-0143624.  
PR 15-JUL-1999; 990S-0144005.  
PR 16-JUL-1999; 990S-0144085.  
PR 16-JUL-1999; 990S-0144086.  
PR 19-JUL-1999; 990S-0144325.  
PR 19-JUL-1999; 990S-0144331.  
PR 19-JUL-1999; 990S-0144332.  
PR 19-JUL-1999; 990S-0144333.  
PR 19-JUL-1999; 990S-0144334.  
PR 19-JUL-1999; 990S-0144335.  
PR 20-JUL-1999; 990S-0144352.  
PR 20-JUL-1999; 990S-0144632.  
PR 20-JUL-1999; 990S-0144884.  
PR 21-JUL-1999; 990S-0144814.  
PR 21-JUL-1999; 990S-0145086.  
PR 21-JUL-1999; 990S-0145088.  
PR 22-JUL-1999; 990S-0145085.  
PR 22-JUL-1999; 990S-0145087.  
PR 22-JUL-1999; 990S-0145089.  
PR 22-JUL-1999; 990S-0145192.  
PR 23-JUL-1999; 990S-0145145.  
PR 23-JUL-1999; 990S-0145218.  
PR 23-JUL-1999; 990S-0145224.  
PR 26-JUL-1999; 990S-0145276.  
PR 27-JUL-1999; 990S-0145913.  
PR 27-JUL-1999; 990S-0145918.  
PR 27-JUL-1999; 990S-0145919.  
PR 28-JUL-1999; 990S-0145951.  
PR 02-AUG-1999; 990S-0146386.  
PR 02-AUG-1999; 990S-0146388.  
PR 02-AUG-1999; 990S-0146389.  
PR 03-AUG-1999; 990S-0147038.  
PR 04-AUG-1999; 990S-0147204.  
PR 04-AUG-1999; 990S-0147302.  
PR 05-AUG-1999; 990S-0147192.  
PR 05-AUG-1999; 990S-0147260.  
PR 06-AUG-1999; 990S-0147303.  
PR 06-AUG-1999; 990S-0147416.  
PR 06-AUG-1999; 990S-0147493.  
PR 09-AUG-1999; 990S-0147935.



PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159350.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 59.68; Score 34; DB 21; Length 155;  
Best Local Similarity 60.08; Pred. No. 60;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
OY 2 PAPMMKFFTT 11  
| | :||||

Db 9 prpshfflt 18  
RESULT 14  
AAG23777  
ID AAG23777 standard; Protein: 159 AA.  
XX AAG23777;  
XX 17-OCT-2000 (first entry)  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 27209.  
DE Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
OS Arabidopsis thaliana.  
XX Ep1033405-A2.  
XX 06-SEP-2000.  
PD 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-0301439.  
PF 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134370.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140921.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.

PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 59 6%; Score 34; DB 21; Length 159;  
Best Local Similarity 60.0%; Pred. No. 62;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
Qy 2 PDPMKFTT 11  
| | :|||  
Db 9 prpshfftt 18

RESULT 15  
AAG32031  
ID AAG32031 standard; Protein; 162 AA.  
XX

AC MAG32031;  
XX 17-OCT-2000 (first entry)  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 38567.  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
OS Arabidopsis thaliana.  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137232.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139889.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144864.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147302.  
PR 04-AUG-1999; 99US-0147304.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147935.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148177.  
PR 11-AUG-1999; 99US-0148313.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.

PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161922.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 59.6%; Score 34; DB 21; Length 162;  
Best Local Similarity 60.0%; Pred. No. 63;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 2 PAPMKFFTT 11  
+ + :||||  
DB 9 prpshriftt 18

Search completed: June 11, 2002, 15:50:48  
Job time: 867 sec